

Db 394 GlnAlaValGlyAlaIleGlyLysGlyAlaHisIleLysGlnLeuAlaArgPhe 413
 Qy 1170 GCCAGCGCTCCATCAAGATTGACACCCAGCCAGCCAGCTGACTCCAAAGTTTCGTATGTT 1229
 Db 414 AlaGlyAlaSerIleLysIleAlaProAlaGluGlyProAspValSerGluArgMetVal 433
 Qy 1230 ATCATCATGAGCCAGAGCCCAATCAAGGCTCAGGGAAGAAATCTATGGCAAACTC 1289
 Db 434 IleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeu 453
 Qy 1290 AAGGAGGAGACTCTTTGGTCCCAAGGAGGAGTGAAGCTGGAGACCCACATACGTGTG 1349
 Db 454 LysGluGluAsnPheAsnProLysGluGluValLysLeuGluAlaHisIleArgVal 473
 Qy 1350 CCAGCATCAGCAGCTGGCGCGGTCTATTGGCAAGGTGGAAACCGTGAACGAGTTGCAG 1409
 Db 474 ProSerSerThrAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGln 493
 Qy 1410 AATTGACGCGAGCTGAGGTGTAGTACCAAGAGACCGACCCCTGATGAGAACACGACG 1469
 Db 494 AsnLeuThrSerAlaGluValIleValProArgAspGlnThrProAspGluAsnGluGlu 513
 Qy 1470 GTCATCGTGAATCATCGGACATTTCTATGCGCTCAGTCCAGTCAACGGAAGATCCGA 1529
 Db 514 ValIleValArgIleGlyHisPheAlaSerGlnThrAlaGlnArgLysIleArg 533
 Qy 1530 GACATCTCGGCCAGGTTAAGCAGCAGCATCATCAAGAGGAGGACAGAGTAACAGCCGACG 1589
 Db 534 GluIleValGlnGlnValLysGlnGlnGlnLysTyrProGlnGlyValAlaSerGln 553
 Qy 1590 CGGAGGAG 1598
 Db 554 ArgSerLys 556
 RESULT 14
 Q7QF9 PRELIMINARY; PRT; 545 AA.
 AC Q7QF9;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=2388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickens M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalek U., Smailus D.B., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC054552; AAH54552.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 545 AA; 59679 MW; 88DD35FF30DF3091 CRC64;
 Alignment Scores:
 Pred. No.: 1.52e-87 Length: 545
 Score: 1299.00 Matches: 257
 Percent Similarity: 74.58% Conservative: 57
 Best Local Similarity: 61.05% Mismatches: 67
 Query Match: 36.79% Indels: 40
 DB: 11 Gaps: 5
 US-09-270-437D-7 (1-1946) x Q7QF9 (1-545)
 Qy 96 GTTTCCTCCGGGTAGGAGCGCGGAAATTCATAATCCGAAATATTCACCCAGCTCCGATGG 155
 Db 75 SerLysLysLeuArgSerArgArgIleGlnIleArgAsnIleProProHisLeuGlnTrp 94
 Qy 156 GAAGTACTGGACAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 215
 Db 95 GluValLeuAspGlyLeuLeuAlaGluTyrGlyThrValGluAsnValGluGlnValAsn 114
 Qy 216 ACCGAGAGTGAAGCGGAGTGGTGAATGTGCTACCTATTCCACCGGAGCAGACGAGCAA 275
 Db 115 ThrAspThrGluThrAlaValValAsnValThrTyrMetThrArgGluGluAlaLysLeu 134
 Qy 276 GCCATCATGAAGCTGAATGGCCAGCAGTGTGAGAACCATGCCCTCAAGGTCTCTACATC 335
 Db 135 AlaIleGluLysLeuSerGlyHisGlnPheGluAspTyrSerPheLysIleSerTyrIle 154
 Qy 336 CCCGATGACGACATAGCA-----CAGGACCTGAGATGGGCGCCGAGGGGGTTTGGC 389
 Db 155 ProAspGluGluValSerSerProSerProHisArgAlaArgGluGlnGlyHisGly 174
 Qy 390 TCTCGGGTCAAGCGCGCCGCTCCTGTGGCAGCGGGGCCCGCCAGCCAGCAGCAG 449
 Db 175 -----ProGlySerSerSerGlnAlaArg 182
 Qy 450 CAAGTGGACATCCCTCTCGGTCTGTCGCCACCCAGTATGTGGTGCCATTATTTGCG 509
 Db 183 GlnIleAspPheProLeuArgIleLeuValProThrGlnPheValGlyAlaIleIleGly 202
 Qy 510 AAGAGGGGGCCACCATCCGCAACATCAAAAGACAGCCAGTCCAGATACAGTCGAT 569
 Db 203 LysGluGlyLeuThrIleLysAsnIleThrLysGlnThrGlnSerArgValAspIleHis 222
 Qy 570 AGGAGGAGAGCAGGTCAGCTGCAAAAGCCATCAGTGTGCTCCTCCACCCCTGAGGCG 629
 Db 223 ArgLysGluAsnSerGlyAlaAlaGluLysProValThrIleHisAlaThrProGluGly 242
 Qy 630 TCTCTCTCGCTGTGAAGATGATCTTGGAGATTATGCAATAAGAGCTTAAGACACCAA 689
 Db 243 ThrSerGluAlaCysArgMetIleLeuGluIleMetGlnLysGluAlaAspGluThrLys 262
 Qy 690 ACGGCTGACGAGTTCCTCGAGATCTGCGGCCCATATAACTTTTAGGGGGCTTCATT 749
 Db 263 LeuAlaGluGluValProLeuLysIleLeuAlaHisAsnGlyPheValGlyArgLeuIle 282
 Qy 750 GGCAGGAGGAGCAGCACTGAGAGAGGTAGAGCAAGATACCGAGACAAATAATCACCATC 809
 Db 283 GlyLysGluGlyArgAsnLeuLysIleGluHisGluThrGlyThrLysIleThrIle 302
 Qy 810 TCTCTGTTCAAGACCTTACCCCTTTACACCTTGAGAGACCATCATCTGTGAAGGGGGCC 869
 Db 303 SerSerLeuGlnAspLeuSerIleTyrAsnProGluArgThrIleThrValArgGlyThr 322
 Qy 870 ATCGAGAAATTGTCAGGGCCGAGCAGGAGAAATAATGAAGAAAGTTTCGGGAGGCTATGAG 929
 Db 323 IleGluAlaCysAlaAsnAlaGluIleGluIleMetLysLysLeuArgGluAlaPheGlu 342

```
Db 432 LysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaAspGlyIleAsp 451
QY 1212 TCCAAAGTTGATGGTATCATCACTGGACCGCCAGAGGCCCAATTCAGGCTCAGGGA 1271
Db 452 AlalysGlnArgMetValIleIleSerGlyProProGluAlaGlnPheLysAlaGlnGly 471
QY 1272 AGAATCTATGCAAACTCAAGGAGAGAACTTCTTGGTCCCAAGGAGGAGTGAAGCTG 1331
Db 472 ArgIlePheGlyLysLeuLysGluGluAsnPheGlyProLysGluGluValLysLeu 491
QY 1332 GAGACCCACATACATGTGTGCCAGCATCAGCAGCTGGCGGGTCAATTTGGCAAGAGTGGAAAA 1391
Db 492 GluAlaHisIleLysValProSerPheAlaAlaGlyArgValIleGlyLysGlyLys 511
QY 1392 ACGGTGAAGAGTTGCAAAATTTGACCGCAGCTGAGTGTGTAGTACCAAGAGACAGCAGC 1451
Db 512 ThrValAsnGluLeuGlnAsnLeuThrSerAlaGluValValProArgAspGlnThr 531
QY 1452 CCTGATCAGAACAGCAGGCTCATCGTGAATAATCATCGACATTTCTATGCCAGTCAGATG 1511
Db 532 ProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaSerGlnLeu 551
QY 1512 GCTCAACGGAGATCCGAGCATCTCTGGCCAGGTTAAGCAGCAGCATCAGAAAGGAGCAG 1571
Db 552 AlaGlnArgLysIleGlnGluIleIleSerGlnValArgGlnGlnGlnProLysPro 571
QY 1572 AGT-----AACAGGCCCGCAGCAGGAGGAG 1598
Db 572 SerAlaAlaGlyProProValAlaArgArgLys 582

RESULT 13
QY6M1 PRELIMINARY; PRT; 556 AA.
ID QY6M1 AC QY6M1
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hepatocellular carcinoma autoantigen.
GN P62.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99207072; PubMed=10190901;
RT Zhang J.Y., Chan E.K., Peng X.X., Tan E.M.;
RT "A novel cytoplasmic protein with RNA-binding motifs is an autoantigen
RT in human hepatocellular carcinoma."
RL J. Exp. Med. 189:1101-1110(1999).
CC 1- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AF057352; AAD31596.1; -
DR GO; GO:0005737; C:cytoplasm; TAS.
DR GO; GO:0003723; F:RNA binding; TAS.
DR GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
DR GO; GO:0009386; P:translational attenuation; TAS.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH type_1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR SMART; PF00076; rrm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH_TYPE 1; 4.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 556 AA; 61842 MW; 1EDEFB100443DDC4 CRC64;
```

Alignment Scores:

```
Pred. No.: 5,24e-117 Length: 556
Score: 1697.00 Matches: 336
Percent Similarity: 79.32% Conservative: 63
Best Local Similarity: 66.80% Mismatches: 80
Query Match: 48.06% Indels: 24
```

```
DB: 4 Gaps: 5
US-09-270-437D-7 (1-1946) x QY6M1 (1-556)
QY 96 GCTTCCCGGGTTAGAGCCGCGAAAATTCAAATCCGAAATATTCACCCCGAGCTCCGATGG 155
Db 76 SerLysLysLeuArgSerArgLysIleGlnIleArgAsnIleProPheLysLeuGlnTrp 95
QY 156 GAAGTACTGACAGCCCTGCTGGCTCAGTATGTGTAGTGAAGAACTGTGACCAAGTGAAC 215
Db 96 GluValLeuAspGlyLeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsn 115
QY 216 ACCGAGAGTGAAGAGTGGTGAATGTCACCTATTCACCCGCGAGCAGACAGGCA 275
Db 116 ThrAspThrGluThrAlaValValAsnValThrTyrAlaThrArgGluAlaLysIle 135
QY 276 GCATCATGAAGCTGAATGGCCACCATGTGAGAACCATGCCCTGGAAGGTCTCTACATC 335
Db 136 AlaMetGluLysLeuSerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIle 155
QY 336 CCGCATGAGCAGATAGCA-----CAGGGACCTGAGAAATGGCGCCGAGGGGGCTTGGC 389
Db 156 ProAspGluGluValSerSerProSerProGlnArgAlaGlnArgGlyAspHisSer 175
QY 390 TCTCGGGGTGAGCCCGCCAGGGCTCACCTGTGCGACGGGGGCCCGACCCAGCAGCAG 449
Db 176 SerArgGluGlnGlyHis-----AlaProGlyGlyThrSerGlnAlaArg 190
QY 450 CAAGTGGACATCCCGCTTGGCTCGCTGGTGGCCACCCAGTATGTGGGTGCCATTTATGGC 509
Db 191 GluIleAspPheProLeuArgIleLeuValProThrGlnPheValGlyAlaIleGly 210
QY 510 AAGGAGGGGCCACCATCCGCAACATCACAACACAGACCCAGTCCAGATAGACGTGCAT 569
Db 211 LysGluGlyLeuThrIleLysAsnIleThrLysGlnThrGlnSerArgValAspIleHis 230
QY 570 AGGAGGAGAACGCGAGTGCAGCTGAAAAGCCATCAGTGTGCACTCCACCCCTGAGGC 629
Db 231 ArgLysGluAsnSerGlyAlaAlaGluLysProValThrIleHisAlaThrProGluGly 250
QY 630 TGCTCTCCCTTGTAGATGATCTTGAGATATTCATTAAGAGGCTTAAGACACCAAA 689
Db 251 ThrSerGluAlaCysArgMetIleLeuGluIleMetGlnLysGluAlaAspGluThrLys 270
QY 690 ACGGCTGACAGGTTCCTTGAAGATCTCTGGCCCATTAATTAATTTGTAGGGCGTTCATT 749
Db 271 LeuAlaGluGluIleProLeuLysIleLeuAlaHisAsnGlyLeuValGlyArgLeuIle 290
QY 750 GCGAAGGAGGACGGACCTGAAGAGGTAGACAGATACCGAGACAAATAATCACCATC 809
Db 291 GlyLysGluGlyArgAsnLeuLysIleGluHisGluThrGlyThrLysIleThrIle 310
QY 810 TCCTCGTTGCAAGACCTTACCTTTCAACCCCTGAGAGGACCATCACTGTGAAGGGGCC 869
Db 311 SerSerLeuGlnAspLeuSerIleTyrAsnProGluArgThrIleThrValLysGlyThr 330
QY 870 ATCGAAGAAATTTGTCAGGGCCGAGCAGGAATAATGAAGAAGTTCGGGAGGCGCTATGAG 929
Db 331 ValGluAlaCysAlaSerAlaGluIleGluIleMetLysLysLeuArgGluAlaPheGlu 350
QY 930 AATGATGTGGCTGCCATGAGCTCTACCTGATCCCTGGCCCTGACACCTGGCTGTAGGT 989
Db 351 AsnAspMetLeuAlaValAsnThrHis-----SerGly 361
QY 990 CTTTTCAGAGCTTCATCCAGCGCAGTCCCGCCCTCCAGCAGCGTTACTGGGGCTGCT 1049
Db 362 TyrPhe-----SerSerLeuTyrProHisGln-----PheGly 373
QY 1050 CCTATAGTCTCTTTATGAGAGGCTCCGAGCAGAGATGGTGCAGGTGTTATTCGCGCC 1109
Db 374 ProPheProHisHisSerTyrProGluGlnGluIleValAsnLeuPheLeProThr 393
QY 1110 CAGGCGAGTGGCGCCATCATCGCAAGAGGGGCGCAGCACATCAACAGCTCTCCCGGTTT 1169
```

01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
Vgl RNA binding protein (Decapentaplegic and Vg-related 1, RNA binding protein).
GN DVIIRBP.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_taxID=7955;
RN [1]
SEQUENCE FROM N.A.
RP Zhang Q., Yaniv K., Oberman F., Wolke U., Git A., Fromer M.,
RA Taylor W., Meyer D., Standart N., Raz E., Visraeli J.K.;
RA "Vgl RP intracellular distribution and evolutionarily conserved
RT expression suggest multiple roles during development.";
RL Mech. Dev. 0:0-0(1999).
RN [2]
SEQUENCE FROM N.A.
RP STRAIN=AB; TISSUE=Body;
RC MEDLINE=22388257; PubMed=12477932;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
SEQUENCE FROM N.A.
RP STRAIN=AB; TISSUE=Body;
RC Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AF161270; RAD45610.1; -;
DR EMBL; BC045873; AAH45873.1; -;
DR ZFIN; ZDB-GENE-000308-1; dvripb.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH type 1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; xrm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS00084; KH type 1; 4.
DR PROSITE; PS0102; RRM; 2; -;
SQ SEQUENCE 582 AA; 63351 MW; 9DAE63200681B306 CRC64;

Alignment Scores:

Pred. No.:	6,96e-127	Length:	582
Score:	1830.50	Matches:	370
Percent Similarity:	82.19%	Conservative:	50
Best Local Similarity:	72.41%	Mismatches:	74
Query Match:	51.84%	Indels:	17
DB:	13	Gaps:	6

US-09-270-437d-7 (1-1946) x Q9PW80 (1-582)

QY 102 CGGGTTAGGACGGGAAATTCGAATCCGAATATTCCACCCCGCTCCGATGGGAAGTA 161

Db	ArgGlnArgSerCysLysLeuGlnIleArgAsnIleProProHisMetGlnTyrGluVal	96
QY	CTGCACAGCCTGCTGGCTCAGTATGTGTACAGTACAGAACTGTGTAGCAAGTGAACACCGAG	221
Db	LeuAspGlyLeuLeuAlaGlnTyrGlyThrValGluSerCysGluGlnValAsnThrAsp	116
QY	AGTCAGACGGCAGTGGTGAATGTACCATATCCACCGGAGGAGCAGACCGCAGCCATC	281
Db	ThrGluThrAlaValValAsnValArgTyrGlyAlaLysAspGlnAlaArgGluAlaMet	136
QY	ATGAAGCTGAATGCCACACCTGTGGAGAACCATCCCTGAAGGTCTCTATCATCCCCGAT	341
Db	AspLysLeuAsnGlyPheLeuMetGluAsnTyrAlaLeuLysValSerTyrIleProAsp	156
QY	GAG---CAGATAGCACAGGGACCTGAGATGGGCGCGAGGGGGCTTTGGCTCTCGGGGT	398
Db	GluThrAlaAlaAspAlaProAlaValGlyGlyArgArgGlyPheAsnProArgGly	176
QY	CAGCCCGCCGAGGGCTCACCTGTGGCAGCGGGGGCCCGCCAGCCAGCAGCAAGTGCAC	458
Db	ProProArgGlnGlySerProSerLeuGlyAlaArgPro-----LysLeuGlnSerAsp	194
QY	ATCCCCCTTCGGCTCCTGGTCCCCACCCAGTATGTGGTGCATTTATTGGCAAGAGGGG	518
Db	ValProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleGlyLysGluGly	214
QY	GCCACCATCCGCACATCACAAAACAGCAGCCAGTCCCAAGATAGAGTCATAGGAGAG	578
Db	AlaThrIleArgAsnIleThrLysGlnThrHisSerLysIleAspIleHisArgLysGlu	234
QY	AACCGAGTGCAGCTGAAAAGCCATCAGTGTGCATCCACCCCTGAGGGTGTCTCTCC	638
Db	AsnAlaGlyAlaAlaGluLysProIleThrValHisSerThrProGluGlyCysSerSer	254
QY	GCTTGTAAAGTATCTTGGAGATATATGATATAAGAGGCTAAAGGACCAACCAACCGCTC	698
Db	AlaCysArgAsnIleMetGluIleMetGlnLysGluAlaIleAspThrLysIleThrGlu	274
QY	GAGGTTCCTTGAAGATCCTGGCCCATTAATTAATCTTTAGGGCGTCTATTGGCAAGAA	758
Db	GluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGlu	294
QY	GCAGCAACCTTGAAGAGTAGCAAGTACCAAGTACCGAGACAAATAACCATCTCTCTGTG	818
Db	GlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeu	314
QY	CAAGACCTTTACCTTTACACCCCTGAGAGGACCATCATCTGTGAAGGGGCCATCAGAA	878
Db	GlnAspLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyThrLeuAspAla	334
QY	TGTTCCAGGGCCGACAGGAATATGAAGAAGTTCGGAGGCGCTATGAGAAATCATGTG	938
Db	CysAlaLysAlaGluGluIleMetLysLysValArgGluSerTyrTyrGluAsnAspVal	354
QY	GCTGCCATG-----AGCTCTCACCTGATCCCTGGCTGAACTGGCTGCTCTAGTGTCT	992
Db	AlaAlaMetHisLeuGlnSerAsnLeuIleProGlyLeuAsnLeuAlaLeuGlyLeu	374
QY	TTTCCAGCTTTCATCCAGC-----GCAGTCCCGCGCTCCACG	1031
Db	PheProGlyAlaAlaSerGlyGlyIleSerProSerValValSerGlyProProGly	394
QY	ACGGTACTTGGGGTGTCTCCCTATAGTCTCTTATTCAGGCTCCCGACGAGATGGTG	1091
Db	AlaGlnAlaGly-----TyrGlnSerPheGlyAlaGlnMetGluSerGluThrVal	411
QY	CAGGTGTTTATCCCGCCAGGCGAGTGGCGGCATCATCGCAAGAAGGGGCGAGCAGCATC	1151
Db	HisLeuPheIleProAlaLeuAlaValGlyAlaIleGlyLysGlnGlyGlnHisIle	431
QY	AAACAGCTCTCCGGTTTGCACGCGCTCCATCAAGATTGCACCCCGAAACACTGCAC	1211


```
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6; TISSUE=Brain, and Olfactory epithelium;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBU databases.
CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AK011689; BAB27779.1; -.
DR EMBL; AB046173; BAB19755.1; -.
DR EMBL; BC045138; AAB45138.1; -.
DR EMBL; BC049082; AAB49082.1; -.
DR MGD; MGI:1890359; Igf2bp3.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH type 1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; xtm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS00084; KH_TYPE_1; 4.
DR PROSITE; PS0102; RRM; 2.
SQ SEQUENCE 579 AA; 63574 MW; CABD9A4355B392B7 CRC64;

Alignment Scores:
Pred. No.: 2,3e-129 Length: 579
Score: 1864.00 Matches: 378
Percent Similarity: 84.71% Conservatives: 54
Best Local Similarity: 74.12% Mismatches: 60
Query Match: 52.79% Indels: 18
DB: 11 Gaps: 10

US-09-270-437D-7 (1-1946) x Q9CPN8 (1-579)
QY 102 CGGGTAGAGCGGCGGAAATTCGAATCCGAATATTCACCCAGCTCCGATGGGAAGTA 161
DB 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluVal 96
QY 162 CTGGACAGCTCTGGCTCAGTATGGTACAGTAGAGAACTGTGAGCAAGTGAACCCGAG 221
DB 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGAGACGGCAGTGTGATGTACCTATTCCACCCGGGAGCAGACCAAGCAAGCCATC 281
DB 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATGAAGCTGAATGGCCACCAAGTGGAGAACCAATGCCCTGAAGTCTCTACATGCCGAT 341
DB 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY 342 GAGCAGATAGCACAG-----GGACCTGAG--AATGGGGCGCGAGGGGGCTTTGGC 389
DB 157 GluThrAlaAlaGlnGlnAsnProSerProGlnLeuArgGlyArgGlyPro---Gly 175
QY 390 TCTCGGGGTGAGCCCGCCAGGCTCAGCTGTGGCAGCGGGGGCCCGCCAGCAGCAGCAG 449
DB 176 GlnArgGlySerSerArgGlnAlaSerPro-----GlySerValSerLysGlnLys 192
QY 450 CAAAGTGACATCCCTTGGCTCTGGTGGCCACCCAGCATATGTGGTGCCATTTATGGC 509
DB 193 ProCysAspLeuProLeuArgLeuValProThrGlnPheValGlyAlaIleIleGly 212
QY 510 AAGAGGGGGCCACCATCCGCACATCAACAACACAGCCAGTCCCAAGATAGACGTGCAT 569
DB 213 LysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHis 232
QY 570 AGGAAGAGAACCGCAGGTGAGTGAAGAAAGCCATCAGTGTGCATCTCCACCCCTGAGGC 629
DB 233 ArgLysGluAsnThrGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGly 252
QY 630 TGCTCTCCGCTTGTGAAGATGATCTTGGAGATATCATTAAGAGCTGAAGACACAA 689
DB 253 ThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLys 272
```

```
QY 690 ACGCTGACGAGGTTCCTCCCTGAAGATCCTGGCCCATATACTTTGTAGGGCTCTCAT 749
DB 273 PheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIle 292
QY 750 GCGAAGGAAGGACGAACTCAAGAAAGGTAGACGAAGATACCGACACAAAATACCATC 809
DB 293 GlyLysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIle 312
QY 810 TCCTCTGTCGAAGACCTTACCCCTTACACCTTGAGAGACCATCCTGTGTAAGGGGCC 869
DB 313 SerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlySer 332
QY 870 ATCGAAGATTGTGACGGGCGCAGCAGCAAGTAATAAGAAAGTTCCGGGAGCCCTATGAG 929
DB 333 ValGluThrCysAlaLysAlaGluGluIleMetLysIleArgGlnSerTyrGlu 352
QY 930 AATGATGTGGTCCCATGAGC-----TCTCACCTGATCCCTGGCTGAACCTGGCTGT 983
DB 353 AsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAla 372
QY 984 GTAGTCTTTTCCGAGCTTCATCCAGCGCAGTCCCGCG-----CTCCACGACG 1034
DB 373 LeuGlyLeuPheProProThrSerGlyMetProProProThrSerGlyProProSerThr 392
QY 1035 GTTACTGGGGTGTCTCCCTATATAGTCTTATGCAAGGCTCCCGACGAGAGATGTGCAG 1094
DB 393 LeuThr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHis 409
QY 1095 GTGTTTATCCCGCCGAGCAGTGGCGCCATCATCGGCAAGAGGGGAGCAGCATCAAA 1154
DB 410 LeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGlnHisLys 429
QY 1155 CAGCTCTCCGGTTGCCAGGCTCCATCAAGATTGCACACCCGAAACACTGACTCC 1214
DB 430 GlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAla 449
QY 1215 AAGTTTCGTATGTTTATCATCTGACCGCCGAGGCGCCAAATTCAGGCTCAGGGAAGA 1274
DB 450 LysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArg 469
QY 1275 ATCTATGGCAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAG 1334
DB 470 IleTyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGlu 489
QY 1335 ACCCATACATGTGTGCCAGCATCAGCAGCTGCGGGGTTCATGGCAAGAGTGGAAAAAG 1394
DB 490 AlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThr 509
QY 1395 GTGAACGAGTTCCAGAATTTGACGCGCAGCTGAGGTGGTAGTACCAAGACACGACCCCT 1454
DB 510 ValAsnGluLeuGlnSerLeuSerSerAlaGluValValProArgAspGlnThrPro 529
QY 1455 GATGAAGACGACCGAGTTCATCGTGAATAATCATCGACATTTCTATGCCAGTCAAGTGGCT 1514
DB 530 AspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCysGlnValAla 549
QY 1515 CAAACGAGATCCGAGACATCTCGCCCGCAGGTTCAGCAG---CAGCATCAGAGGGA--- 1568
DB 550 GlnArgLysIleGlnGluLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeu 569
QY 1569 CAGAGTAAACCGCCGCGCAGCAGGAGGAAG 1598
DB 570 GlnSerGlyProProGlnSerArgLys 579

RESULT 11
Q8C2J9 PRELIMINARY; PRT; 579 AA.
AC Q8C2J9
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Insulin-like growth factor 2.
GN IGFBP3.
```

QY 513 GAGGGGCGCCATCCGCAACATCACAAACAGACCAGTCCAGATAGAGTGCATAGG 572
Db 214 GluGlyAlaThrIleArgAsnIleThrIleGlnThrGlnSerIleAspValHisArg 233
QY 573 AAGGAGAACGCGAGTGGAGCTGAAAGACCATCATGTGCACTCCACCCCTGAGGGTGC 632
Db 234 LysGluAsnAlaGlyAlaAlaGluGlySerIleThrIleLeuSerThrProGluGlyThr 253
QY 633 TCCTCCGCTGTTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGACACCAAAACG 692
Db 254 SerAlaAlaCysLysSerIleLeuGluIleMethIleLysGluAlaGlnAspIleLysPhe 273
QY 693 GCTGACAGGTTCCCTGAGATCCTGGCCCATTAATTAACCTTTAGGGCGTCTCATGGC 752
Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY 753 AAGGAGGACGGAACCTGAAGAGTAGACGACATACCGAGACAAATAACCATCTCC 812
Db 294 LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
QY 813 TCGTTGCAAGACCTTACCTTTACACCCCTGAGAGCCATCACTGTGAAGGGGCCATC 872
Db 314 ProLeuGlnLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY 873 GAGAAATTGTCAGGGCGGACGAGAAATATGAAAGTTCGGGAGGCCCTATGAGAT 932
Db 334 GluThrCysAlaLysAlaGluGluGluIleMethLysIleArgGluSerTyrGluAsn 353
QY 933 GATGTGCTGCATAGC-----TCNCACTGATCCCTGCGCCGACCTGGCTGCTGTA 986
Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaGlu 373
QY 987 GGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCTCCAGCAGCGTT 1037
Db 374 GlyLeuPheProProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
QY 1038 ACTGGGCTGCTCCTTAGCTCTTATGAGCTCCGAGAGCTCCGAGAGAGATGTCAGGTG 1097
Db 394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisGln 410
QY 1098 TTTATCCCGCCAGGAGTGGGGCCATCATCGCAAGAGGAGGAGGAGCAGCATCAACAG 1157
Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlySerGlnGlyGlnHisIleLysGln 430
QY 1158 CTCTCCCGTTTGCAGCGCTTCATCAAGATTGCACACCCGCAAAACCTGACTCCAAA 1217
Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY 1218 GTTCGTATGTTATCATCATCTGAGCCCGCAGAGGCCCAATTCAGGCTCAGGAGATC 1277
Db 451 ValArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY 1278 TATGGCAACTCAAGAGGAGAGACTCTTTGGTCCCAAGGAGGAGTGAAGCTGGAGCC 1337
Db 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY 1338 CACATACGTGTGCAGCATCAGCAGCTGGCGGGTCAATGCAAGGTGGAAAAACGGTG 1397
Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
QY 1398 AACGAGTGCAGAAATTCAGCGGAGCTGAGGTGTAGTACCAAGAGCAGACCCCTGAT 1457
Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
QY 1458 GAGAACGACGAGTTCATCGTGAATAATCATCGACATTTCTATGCGAGTCCAGATGGCTCAA 1517
Db 531 GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
QY 1518 CGGAAGATCCGAGACATCTCTGGCCCGCAGGTAAAGCAG---CAGCATCAGAAGGGA---CAG 1571
Db 551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnLysAlaLeuGln 570
QY 1572 AGTAACCGAGCCCGCAGGACGAGGAAG 1598

Db 571 SerGlyProGlnSerArgLys 579
RESULT 10
Q9CPN8 PRELIMINARY; PRT; 579 AA.
ID Q9CPN8;
AC Q9CPN8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 10 days embryo cDNA, RIKEN full-length enriched library,
DE clone:2610035818, full insert sequence (Igf2 mRNA-binding protein 3)
DE (Insulin-like growth factor 2, binding protein 3).
GN IGFBP3 OR 2610101N1RIK OR MIMP3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Embryo;
RC MEDLINE=21085660; PubMed=11217851;
RX Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Arakawa K., Iizawa M., Nishi K., Kiyosawa H., Kondo S., Yamakawa I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirral L.M., Staehli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA Hayashizaki Y.
RT "Functional annotation of a full-length mouse cDNA collection.";
EL Nature 409:685-690(2001).
RN [2]
SEQUENCE FROM N.A.
RP Mori H., Sakakibara S., Imai T., Nakamura Y., Iijima T., Suzuki A.,
RA Yuasa Y., Takeda M., Okano H.;
RT "Expression of mouse Igf2 mRNA-binding protein 3 and its implications
RT for the developing central nervous system.";
RL J. Neurosci. Res. 0:0-0(2001).
RN [3]
SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Brain, and Olfactory epithelium;
RC MEDLINE=22388257; PubMed=12477932;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Strausberg R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uesdin T.B., Toehiyuki S., Carninci P., Prange C.,
RA Raha S.S., Ioquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madao A., Rodriguez S., Sanchez A.,
RA Whiting M., Madao A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

Db 254 ThrIleHisSerThrProGluGlyCysSerAlaAlaCysLysLysIleIleMetGluIleMet 273
Qy 666 CATAAAGAGGCTAAGGACACCAAAACGGCTACGAGGTTCCCTCGAAGATCTCGCCCAT 725
Db 274 GlnLysGluAlaGlnAspThrLysPheThrGluGluIleProLeuLysIleLeuAlaHis 293
Qy 726 AATACTTTGATGGCGCTCATTTGGCAAGGACGGAACCTGAAGGTAGAGCAA 785
Db 294 AsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIleGluGln 313
Qy 786 GATACCGAGACAAAATACCATCTCTCGTTGCAAGACCTTACCTTTACAACCTCGAG 845
Db 314 AspThrAspThrLysIleThrIleSerProLeuGlnAspLeuThrLeuTyAsnProGlu 333
Qy 846 AGGACCATCACTGTGAAGGGCCCATCGAATTTGTCAGGCGCCGACGAGAAATATG 905
Db 334 ArgThrIleThrValLysGlySerIleGluThrCysAlaLysAlaGluGluValMet 353
Qy 906 AAGAAAGTTCGGAGCGCTATGAGATGATGGCTGCCATGAGC-----TCTCACCTG 959
Db 354 LysLysIleArgGluSerTyGluAsnAspIleAlaAlaMetAsnLeuGlnAlaHisLeu 373
Qy 960 ATCCCTGGCTGAACCTGGCTCTGCTAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCG 1019
Db 374 IleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPheProProSerSerSerGlyMetPro 393
Qy 1020 CGCCTCCCT-----AGCAGCCTTACTGGGGTGCTCCCTATAGTCTCTTTATGCGAG 1070
Db 394 ProProSerAlaGlyValSerSerProThrThrSerAlaSerTyProProPheGlyGln 413
Qy 1071 GTCCTCGAGCAGGAGATGTGTGAGTGTATTCCTCCGCGCCAGCGAGTGGCGCCATCATC 1130
Db 414 GlnProGluSerGluThrValHisLeuPheIleProAlaLeuAlaValGlyAlaIleIle 433
Qy 1131 GCGAAGAGGGCGAGCACATCAAAAGCTCTCCCGTTTGCCAGCGCTCCATCAAGATT 1190
Db 434 GlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIle 453
Qy 1191 GCACCAACCGAAACCTGACTCCAAAGTTGTATGTTATCATCTGAGCGCCGAGAG 1250
Db 454 AlaProAlaGluGlyProAspAlaLysLeuArgMetValIleIleThrGlyProProGlu 473
Qy 1251 GCGCAATTCAGGCTCAGGAGAGATCTATGCAACCTGATGCAACCTCAGGAGAGACTTCTTGT 1310
Db 474 AlaGlnPheLysAlaGlnGlyArgIleTyGlyLysLeuLysGluGluAsnPheGly 493
Qy 1311 CCCAAGGAGAGTGAAGCTGAGAGCCACATACCTGTGCCAGCATCAGCAGTGGCCGG 1370
Db 494 ProLysGluGluValLysLeuGluAlaHisIleLysValProSerTyAlaAlaGlyArg 513
Qy 1371 GTCATTGGCAAGGTGGAAAAACGGTGAACGAGTTTCAGAAATTCAGCGCAGCTGAGTG 1430
Db 514 ValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThrSerAlaGluVal 533
Qy 1431 GTAGTACCAAGAGACAGACCCCTGATGAGACGACAGCTCATCTGTAATCATCGGA 1490
Db 534 ValValProArgAspGlnThrProAspGluAsnAspGlnValValLysIleThrGly 553
Qy 1491 CATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCTGCGCCAGGTTAAG 1550
Db 554 HisPheTyAlaSerGlnLeuAlaGlnArgLysIleGlnIleLeuAlaGlnValArg 573
Qy 1551 CAGCAGCATCAGAAG-----GGACAGAGTAACAGGCCAGGACCGAGGAG 1598
Db 574 ArgGlnGlnGlnGlnGlnLysThrAlaGlnSerGlyGlnProGlnProArgArgLys 593
RESULT 9
ID 000425 PRELIMINARY; PRT; 579 AA.
AC 000425;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Putative RNA binding protein KOC (KOC).
GN KOC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Mueller-Pillasch F., Lacher U., Wallrapp C., Et AL.;
RL Oncogene 0:0-0(0).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA Mueller-Pillasch F., Lacher U., Wallrapp C.;
RL Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.
CC - SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; U97188; AAC35208.1; -.
DR EMBL; U76705; AAD09223.1; -.
DR GO; GO:0005737; C:cytoplasm; TAS.
DR GO; GO:0003723; F:RNA binding; TAS.
DR GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
DR GO; GO:0006412; P:protein biosynthesis; TAS.
DR GO; GO:0006396; P:RNA processing; TAS.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH_type_1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS00084; KH_TYPE_1; 4.
DR PROSITE; PS0102; RRM; 2.
DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
SQ SEQUENCE 579 AA; 63720 MW; AESC3A8EE3C135C5 CRC64;
Alignment Scores:
Pred. No.: 4, 96e-130 Length: 579
Score: 1873.00 Matches: 380
Percent Similarity: 84.48% Conservative: 50
Best Local Similarity: 74.66% Mismatches: 63
Query Match: 53.04% Indels: 16
DB: 4 Gaps: 9
US-09-270-437D-7 (1-1946) x 000425 (1-579)
Qy 102 CGGGTTAGAGCGCGGAAATTCAAATTCGAAATATTCACCCACGCTCGATGGAGTA 161
Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluVal 96
Qy 162 CTGGACAGCGCTCGCTCAGTATGTACAGTAGAGAACTGTGAGCAAGTGAACACCGAG 221
Db 97 LeuAspSerLeuLeuValGlnTyGlyValValGluSerCysGluGlnValAsnThrAsp 116
Qy 222 AGTAGACGCGCAGTGTGAATGTCACTATTCCAAACCGGAGCAGACCGCAAGCCATC 281
Db 117 SerGluThrAlaValValAsnValThrTySerSerLysAspGlnAlaArgGlnAlaLeu 136
Qy 282 ATGAAGCTGAATGGCCACCAGTTGAGAACCATCCCTGAAAGTCTCTACATCCCGAT 341
Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyIleProAsp 156
Qy 342 GAGCAGATAGCA--CAGGACCTCAGAATGGCGCCGAGGG-----GGCTTTGGCTCT 392
Db 157 GluMetAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGln 176
Qy 393 CGGGGTTCAGCCCGCCAGCGGCTCACGTGTGGCAGCGGGGGCCCCAGCCAGCAGCAA 452
Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
Qy 453 GTGGACATCCCTTCGGCTCTGCTGTCGCCACCCAGTATGTGGTCCCATTTATGGCAG 512
Db 194 CysAspLeuProLeuArgLeuValProThrGlnPheValGlyAlaIleLeuGlyLys 213

Db 359 SerLeuGlnSerHisLeuLeuProGlyLeuAsnLeuAlaAlaValGlyLeuPheProAla 378
QY 1002 TCATCCAGCGAGTCCCGCCCTCCAGCAGCGTTACTGGGCTGCTCCCTATAGTCC 1061
Db 379 SerSerAlaValProProProProSerSerValThrGlyAlaAlaProTy-SerSer 398
QY 1062 TTTATGCGAGTCCCGCAGGAGATGGTGAGGTGTTATCCCGCCAGCGAGTGGC 1121
Db 399 PheMetGlnAlaProGluGlnGluMetValGlnValPheIleProAlaGlnAlaValGly 418
QY 1122 GCCATCATCGCAAGAGGGGCGACACATCAACAGCTCTCCCGTTGGCCAGCGCTCC 1181
Db 419 AlaIleIleGlyLysLysGlyGlnHisIleLysGlnLeuSerArgPheAlaSerAlaSer 438
QY 1182 ATCAAGATTGCAACCCCGCAACACCTGACTCCAAAGTTCGTATGTTATCATCATCGGA 1241
Db 439 IleLysIleAlaProProGluThrProAspSerLysValArgMetValIleThrGly 458
QY 1242 CCGCCAGAGCCCAATTCAGGGCTCAGGAGAGATCTATGCAACTCAAGGAGGAGAC 1301
Db 459 ProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrglyLysLeuLysGluGluAsn 478
QY 1302 TTCTTTGTCCTCCAGGAGGAAGTGAAGCTGAGACCCACATACCTGTGCGCAGCATCAGCA 1361
Db 479 PhePheGlyProLysGluGluValLysLeuGluThrHisIleArgValProAlaSerAla 498
QY 1362 GCTGGCGGGTCTATGTCGCAAGTGGGAAAACCGGTGAACGAGTTGCGAGAATTTGACGGCA 1421
Db 499 AlaGlyArgValIleGlyLysGlyGlyLysThrValAsnGluLeuGlnAsnLeuThrAla 518
QY 1422 GCTGAGGTGTAGTACCAAGAGACGAGACCCCTGATGAGACGACGAGTTCATCGTAAA 1481
Db 519 AlaGluValValProArgAspGlnThrProAspGluAsnAspGlnValIleValLys 538
QY 1482 ATCATCGCAATTTCTATGCGAGTCAGATGGCTCAACGGAAGATCCGAGACATCCTGGCC 1541
Db 539 IleIleGlyHisPheTyAlaSerGlnMetAlaGlnArgLysIleArgAspIleLeuAla 558
QY 1542 CAGTTTAAGCAGCAGCATGAGAGGGACAGAGTAACACGAGCCCGCAGGACGAGGAAG 1598
Db 559 GlnValLysGlnGlnHisGlnLysGlyGlnSerAsnLeuAlaGlnAlaArgArgLys 577

RESULT 6

O42254 PRELIMINARY; PRT; 576 AA.
AC O42254;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Zipcode-binding protein.
GN ZBP1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]_TaxID=9031;
RP SEQUENCE FROM N.A.
RX MEDLINE=9722007; PubMed=9121465;
RA Ross A.F., Olevnikov Y.S., Kislauskis E.H., Taneja K.L., Singer R.H.;
RT "Characterization of a beta-actin mRNA zipcode-binding protein.";
RL Mol. Cell. Biol. 17:2158-2165 (1997).
CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AF026527; AAB82295.1; -.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH type 1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH_TYPE_1; 4.

DR PROSITE; PS50102; RRM; 2.
DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
SQ SEQUENCE 576 AA; 63271 MW; 01AAAF2D1D81C8811 CRC64;
Alignment Scores:
Pred. No.: 2.87e-168 Length: 576
Score: 2389.50 Matches: 471
Percent Similarity: 96.59% Conservative: 11
Best Local Similarity: 94.39% Mismatches: 14
Query Match: 67.67% Indels: 3
DB: 13 Gaps: 2
US-09-270-437D-7 (1-1946) x O42254 (1-576)

QY 108 AGGAGCCGGAATAATCAAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTACTGAC 167
Db 79 ArgSerArgLysIleGlnIleArgAsnIleProGlnLeuArgTrpGluValLeuAsp 98
QY 168 AGCTGCTGCTCAGTATGGTACAGTAGAGAACTGTGAGCAAGTGAACACCGAGAGTGA 227
Db 99 GlyLeuLeuAlaGlnTyrglyThrValGluAsnGlyGlnValAsnThrAspSerGlu 118
QY 228 ACGGCAGTGTGATGTACCTATTCACCGGGAGCAGACCGGAGCCATCATGAG 287
Db 119 ThrAlaValValAsnValThrTyThrAsnArgGluGlnThrArgGlnAlaIleMetLys 138
QY 288 CTGAATGCCACACAGTGTGGAGAACCATGCTGAAGGTCTCTATATCCCGATGAGCAG 347
Db 139 LeuAsnGlyHisGlnLeuGluAsnHisValSerTyrlleProAspGluGln 158
QY 348 ATAGCAGAGGACCTGAGATGGCGCGGAGGGGCTTTGGCTTCGGGGTACGCCCGC 407
Db 159 SerValGlnGlyProGluAsnGlyArgArgGlyGlyPheGlyAlaArgGlyAlaProArg 178
QY 408 CAGGGCTCAGCTGTGGCAGCGGGGCGCCAGCCAGCAGCAGCAGTGGACATCCCTT 467
Db 179 GlnGlySerProValThrAlaGlyAlaProValLysGlnProValAspIleProLeu 198
QY 468 CGGCTCCTGTGTCGCCACCCAGATATGTGGTGCCTATTTGGCAAGGAGGGGCCACCATC 527
Db 199 ArgLeuLeuValProThrGlnTyValGlyAlaIleIleGlyLysGluGlyAlaThrIle 218
QY 528 CGCAACATCAAAACAGACCCAGTCCAAAGATGACGTGATAGAGAGGAGAACCGAGGT 587
Db 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGly 238
QY 588 GCAGCTCAAAAGCCATCAGTGTGCATCCACCCCTGAGGGCTGCTCCTCGCTTGAAG 647
Db 239 AlaAlaGluLysAlaIleSerIleHisSerThrProGluGlyCysSerAlaAlaCysLys 258
QY 648 ATGATCTTGAGATTATGCATAAAGAGGCTTAAGGACACCAAAACCGCTGACGAGTTCCC 707
Db 259 MetIleLeuGluIleMetGlnLysGluAlaLysAspThrLysThrAlaAspGluValPro 278
QY 708 CTGAAGATCTGGCCCATATAACTTTGTAGGGCGTCTCATTTGCAAGAGGACGGAAC 767
Db 279 LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn 298
QY 768 CTGAAGAAGGTAGCAAGATACCGAGACAAAATCACCATCTCTCTGTTGCAAGACTT 827
Db 299 LeuLysLysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeu 318
QY 828 ACCCTTTACAACCCCTGAGAGGACCATCTGTGAAGGGGGCCATCGAGATTTGTTGAGG 887
Db 319 ThrLeuTyAsnProGluArgThrIleThrValLysGlySerIleGluAsnGlyLys 338
QY 888 GCGGACGAGAAATATGAAGAAAGTTCGGGAGCGCTATGAGATGATGTGGTGCATG 947
Db 339 AlaGluGlnGluIleMetLysLysValArgGluAlaTyrgluAsnAspValAlaAlaMet 358
QY 948 AGC-----TCTCACCTGATCCCTGGCTGMACTGGCTGCTGTAGTCTTTTCCAGGCT 1001
Db 359 SerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaAlaValGlyLeuPheProAla 378

```
QY 1242 CGCCAGAGGCCCAATTCAGGCTCAGGGAAGAATCTATGGCAAACTCAGAGGAGAAC 1301
|||
Db 459 ProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysLeuLysGluGluAsn 478
|||
QY 1302 TTCCTTTGGTCCCAAGGAGGAAGCTGAAGCTGGAGAGCCACATACGTTGTGCAGCATCAGCA 1361
|||
Db 479 PhePheGlyProLysGluGluValLysLeuGluThrHisIleArgValProLaserAla 498
|||
QY 1362 GCTGCCCGGTCATTTGGCAAAAGTGGAAAAACCGTGAACAGTTGCAGAAATTTGACGGCA 1421
|||
Db 499 AlaGlyArgValIleGlyLysGlyGlyLysThrValAsnGluLeuGluAsnLeuThrAla 518
|||
QY 1422 CTTGAGGTGTTAGTACCAAGACACACAGCCCTGATGAGAACGACCAAGTCTATCGTAAA 1481
|||
Db 519 AlaGluValValProArgAspGlnThrProAspGluAsnAspGlnValIleValLys 538
|||
QY 1482 ATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCTTGGCC 1541
|||
Db 539 IleIleGlyHisPheTyrAlaSerGlnMetAlaGlnArgLysIleArgAspIleLeuAla 558
|||
QY 1542 CAGGTTAAGCAGCAGCATCAGAAGGACAGAGTAACACGAGCCAGGAGGAG 1598
|||
Db 559 GlnValLysGlnGlnHisGlnLysGlyGlnSerAsnLeuAlaGlnAlaArgArgLys 577
|||
RESULT 5
Q80US9 PRELIMINARY; PRT; 577 AA.
AC Q80US9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Igfbp1 protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv x 129/Sv-CP; TISSUE=Embryonic stem cells;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv x 129/Sv-CP; TISSUE=Embryonic stem cells;
RA Strausberg R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC051679; AAH51679.1; -.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH_dom.
DR InterPro; IPR004086; KH_type_1.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; xrm; 2.
DR SMART; SM00322; KH; 4.
```

```
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH_TYPE_1; 4.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 577 AA; 63378 MW; D439F544257DA3CE CRC64;

Alignment Scores:
Pred. No.: 2.9e-175 Length: 577
Score: 2484.00 Matches: 492
Percent Similarity: 98.80% Conservative: 1
Best Local Similarity: 98.60% Mismatches: 4
Query Match: 70.35% Indels: 2
DB: 11 Gaps: 1

US-09-270-437D-7 (1-1946) x Q80US9 (1-577)
QY 108 AGAGCCCGGAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTACTGCGAC 167
|||
Db 79 ArgSerArgLysIleGlnIleArgAsnIleProGlnLeuArgTyrGluValLeuAsp 98
|||
QY 168 AGCCTGCTGGCTCAGTATGGTACAGTAGAAGTGTGAGCAAGTGAACACACGAGAGTGG 227
|||
Db 99 SerLeuLeuAlaGlnTyrGlyThrValGluAsnCysGluGlnValAsnThrGluSerGlu 118
|||
QY 228 ACGGAGTGGTCAATGTCACTATTCCAAACCGGAGCAGACCAGGCAAGCCATCATGAAG 287
|||
Db 119 ThrAlaValValAsnValThrTyrSerAsnArgGluGlnThrArgGlnAlaIleMetLys 138
|||
QY 288 CTGAATGGCCACCATGTTGGAGAACCATGCCCTGAAAGTCTCTCATCCCCGATGAGCAG 347
|||
Db 139 LeuAsnGlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyrIleProAspGluGln 158
|||
QY 348 ATAGCACAGGACCTGAGATGGGCGCGAGGGGCTTTGGCTCTCGGGTCTCAGCCCGC 407
|||
Db 159 IleThrGlnGlyProGluAsnGlyArgArgGlyGlyPheGlySerArgGlyGlnProArg 178
|||
QY 408 CAGGGCTCACCTGTGGCAGCGGGGCGCCAGCAAGCAGCAGCAAGTGGACATCCCCCTT 467
|||
Db 179 GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnProValAspIleProLeu 198
|||
QY 468 CGGCTCTGGTGCCCAACCCAGTATGTGGTGGCCATATTGGCAAGAGGGGCGCCACCATC 527
|||
Db 199 ArgLeuLeuValProThrGlnTyrValGlyAlaIleIleGlyLysGluGlyAlaThrIle 218
|||
QY 528 CGCAATCATCAAAACAGACCCAGTCCAGATAGACGTGCATAGGAGGAGAACGAGT 587
|||
Db 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGlnAsnAlaGly 238
|||
QY 588 GCAGCTGAAAAAGCCATCAGTGTGCATCTCCACCCCTGAGGGTGTCTCTCCGCTTGTAAAG 647
|||
Db 239 AlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCysSerSerAlaCysLys 258
|||
QY 648 ATGATCTTGGAGATATGATTAAGAGGCTAAGGACACCAAAACCGCTGACAGGTTGCC 707
|||
Db 259 MetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThrAlaAspGlyValPro 278
|||
QY 708 CTCAGATCTCGGCCCATTAATTAATCTGTAGGGCTCTCATTTGGCAAGAGGAGCGAAC 767
|||
Db 279 LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn 298
|||
QY 768 CTGAAGAGGTAGAGCAAGATACCGAGACAAAATATCACCATCTCTCTGTTGCAAGACCTT 827
|||
Db 299 LeuLysLysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeu 318
|||
QY 828 ACCCTTTACACCTGAGAGGACCAATCACTGTGAAGGGGCCCATCAGAAATTTGTCAGG 887
|||
Db 319 ThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIleGluAsnCysCysArg 338
|||
QY 888 GCGGAGCAGGAATAATGAGAGAACTTCGGAGGCTTATGAGATCATCATGCTGCTGCATG 947
|||
Db 339 AlaGluGlnGluIleMetLysLysValArgGluAlaTyrGluAsnAspValAlaAlaMet 358
|||
QY 948 AGC-----TCTCACCTGATCCCTGGCTGAAACCTGGCTGCTGTAGGTCTTTTCCAGCT 1001
|||
```



```

QY 1242 CCGCCAGAGCCCAATTCAGGCTCAGGAGAACTCTATGCAAACTCAAGGAGGAAC 1301
Db |||||
QY 459 ProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysLeuLysGluGluAsn 478
Db |||||
QY 1302 TTCCTTGGTCCCAAGGAGGAAGTCAAGCTGGAGACCCACATACCTGTGCCAGCATCAGCA 1361
Db |||||
QY 479 PhePheGlyProLysGluGluValLysLeuGluThrHisIleArgValProAlaSerAla 498
QY 1362 GCTCGCGGGTCATTGGCAAGGTGGAAAAACGGTGAAACGGTGAACGGTTCAGCAATTCACGGCA 1421
Db |||||
QY 499 AlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThrAla 518
QY 1422 GCTCAGGTGTAGTACCAAGAGACGACACCCCTGATGAGACGACCCAGGTCTATCGTGA 1481
Db |||||
QY 519 AlaGluValValProArgAspGlnThrProAspGluAsnAspGlnValIleValLys 538
QY 1482 ATCATCCGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCTGGCC 1541
Db |||||
QY 539 IleIleGlyHisPheTyrAlaSerGlnMetAlaGlnArgLysIleArgAspIleLeuAla 558
QY 1542 CAGTTAAGCAGCAGCATCAGAACGGGACAGAGTAAACGAGGCCCGGACCGGAGGAAG 1598
Db |||||
QY 559 GlnValLysGlnGlnHisGlnLysGlyGlnSerAsnLeuAlaGlnAlaArgArgLys 577

```

RESULT 4

```

Q8BRH1 ID Q8BRH1 PRELIMINARY; PRT; 577 AA.
AC Q8BRH1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Insulin-like growth factor 2.
GN IGF2BP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK044850; BAC32119.1; -.
DR MGD; MGI:1890357; Igf2bp1.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH type_1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR SMART; SM00360; RRM; 2.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PSS0084; KH TYPE_1; 4.
DR PROSITE; PSS0102; RRM; 2.
SQ SEQUENCE 577 AA; 63378 MW; D75897450841532E CRC64;

```

Alignment Scores:

```

Pred. No.: 2.9e-175 Length: 577
Score: 2484.00 Matches: 492
Percent Similarity: 98.80% Conservative: 1
Best Local Similarity: 98.60% Mismatches: 4
Query Match: 70.35% Indels: 2
DB: 11 Gaps: 1

```

US-09-270-437D-7 (1-1946) x Q8BRH1 (1-577)

```

QY 108 AGGAGCCGGAATAATCAAAATCCGAATATTCACCCAGCTCCGATGGGAGTACTGGAC 167
Db |||||
QY 79 ArgSerArgLysIleGlnIleArgAsnIleProGlnLeuArgTrpGluValLeuAsp 98
Db |||||

```

```

QY 168 AGCTCTGCTGCTCAGTATGGTACAGTAGAAGAACTGTGAGCAAGTGAACACCGAGAGTGAG 227
Db |||||
QY 99 SerLeuLeuAlaGlnTyrGlyThrValGluAsnCysGluGlnValAsnThrGluSerGlu 118
QY 228 ACGCAGTGGTGAATGTACACTATTCCAACCGGAGCAGACCCAGCAAGCCATCATGAAG 287
Db |||||
QY 119 ThrAlaValValAsnValThrTyrSerAsnArgGluGlnThrArgGlnAlaIleMetLys 138
QY 288 CTGAATGGCCACCAAGTTGGAGAACCATGCCCCGAGGTCTCTACATCCCGATGAGCAG 347
Db |||||
QY 139 LeuAsnGlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyrIleProAspGluGln 158
QY 348 ATACACAGGAGCACTGAGATGGCGCGGAGGGGCTTTGGCTCTCGGGGTACGCCCCGC 407
Db |||||
QY 159 IleThrGlnGlyProGluAsnGlyArgGlyGlyPheGlySerArgGlyGlnProArg 178
QY 408 CAGGGCTCACCTGTGGCAGCGGGGCCCCAGCCAGCAGCAGCAAGTGGACATCCCCCTT 467
Db |||||
QY 179 GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnProValAspIleProL 198
QY 468 CGGCTCTGTGGTCCCCACCCAGTATGTGGGTGCCATTATTGGCAAGGAGGGGCCACCATC 527
Db |||||
QY 199 ArgLeuLeuValProThrGlnTyrValGlyAlaIleIleGlyLysGluGlyAlaThrIle 218
QY 528 CGCAACATCACAAACAGACCCAGTCCCAAGATAGACGTGCATAGAGAGGAGAACCGAGGT 587
Db |||||
QY 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGly 238
QY 588 GCAGCTCAAAAAGCCATCAGTGTGCACCTCCACCCCTGAGGGCTGCTCTCTCGCTTGTAA 647
Db |||||
QY 239 AlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCysSerSerAlaCysLys 258
QY 648 ATGATCTTGGAGATTATGCATAAAGAGGCTAAGACACCAAAACGGCTGACGAGGTCCCC 707
Db |||||
QY 259 MetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThrAlaAspGluValPro 278
QY 708 CTGAAGATCTGGCCCATATAACTTTGTAGGGGTCTCTCATTTGGCAGGAGGAGCGAAGC 767
Db |||||
QY 279 LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn 298
QY 768 CTGAAGAGGTAGAGCAAGATACCGAGCAAAAATCCACATCTCTCTGTTCCAGACCTT 827
Db |||||
QY 299 LeuLysLysValGluGlnAspThrGluThrLysIleThrIleThrIleThrIleThrIle 318
QY 828 ACCCTTTACAAACCTCGAGAGACCATCACTGTGAAGGGGGCCATCGAGAATTTGTTCAGG 887
Db |||||
QY 319 ThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIleGluAsnCysCysArg 338
QY 888 GCGGAGCAGGAATATGAAGAAAGTTCCGGAGGCCCTATGAGATGATGTGGCTGCCATG 947
Db |||||
QY 339 AlaGluGlnGluIleMetLysLysValArgGluAlaTyrGluAsnAspValAlaAlaMet 358
QY 948 AGC-----TCTCACTGATCCCTGGCTGAAACCTGGCTGTGTAGTGTCTTTTCCAGCT 1001
Db |||||
QY 359 SerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaValGlyLeuPheProAla 378
QY 1002 TCATCCAGCGCAGTCCCGCCGCTCCAGAGCCCTTACTGGGGTGTCTCCCTATAGTCTCC 1061
Db |||||
QY 379 SerSerSerAlaValProProProProProProProProProProProProProProPro 398
QY 1062 TTTATGAGGCTCCCGAGCAGGAGATGTGTGAGGTGTTTATCCCGCCCGCAGGCGAGTGGGC 1121
Db |||||
QY 399 PheMetGlnAlaProGluGlnGlyMetValGlnValPheIleProAlaGlnAlaValGly 418
QY 1122 GCCATCATCGCAAGAGGGGCGAGCAGCATCAAAACAGCTCTCCCGGTTTTCAGGCGCTCC 1181
Db |||||
QY 419 AlaIleIleGlyLysLysGlyGlnHisIleLysGlnLeuSerArgPheAlaSerAlaSer 438
QY 1182 ATCAAGATTGCACACCCGAAACACCTGACTCCAAAGTTCGTATGTTATCATCACTGGAC 1241
Db |||||
QY 439 IleLysIleAlaProProGluThrProAspSerLysValArgMetValValIleThrGly 458

```

RP SEQUENCE FROM N.A.
 RX MEDLINE=94179348; PubMed=81322663;
 RA Prokipcak R.D., Herrick D.J., Ross J.;
 RT "Purification and properties of a protein that binds to the C-terminal
 RT coding region of human c-myc mRNA";
 RL J. Biol. Chem. 269:9261-9269 (1994).
 RN [4]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=97322234; PubMed=9178888;
 RA Leeds P., Kren B.T., Boylan J.M., Betz N.A., Steer C.J.,
 RA Gruppuso P.A., Ross J.;
 RT "Developmental regulation of CRD-BP, an RNA-binding protein that
 RT stabilizes c-myc mRNA in vitro";
 RL Oncogene 14:1279-1286 (1997).
 RN [5]
 RN SEQUENCE FROM N.A.
 RA Ross J., Prokipcak R.D., Leeds P., Doyle G.A.R., Betz N.A.,
 RA Fleisig A.J.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustingich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690 (2001).
 CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
 DR EMBL; AF061563; AAC72743.1; -;
 DR EMBL; AK013940; BAB29071.1; -;
 DR HSSP; P11940; 1CVJ.
 DR MGD; MGI:1890357; Igf2bp1.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR004087; KH dom.
 DR InterPro; IPR004088; KH_type_1.
 DR Pfam; PF00013; KH; 4.
 DR Pfam; PF00076; rtm; 2.
 DR SMART; SM00322; KH; 4.
 DR SMART; SM00360; RM; 2.
 DR PROSITE; PSS0084; KH_TYPE_1; 4.
 DR PROSITE; PSS0102; RRM; 2.
 SQ SEQUENCE 577 AA; 63450 MW; EFBB1AF2FF9F0344 CRC64;

 Alignment Scores:
 Pred. No.: 8,79e-176 Length: 577
 Score: 2491.00 Matches: 493
 Percent Similarity: 99.00% Conservative: 1
 Best Local Similarity: 98.80% Mismatches: 3
 Query Match: 70.55% Indels: 2
 DB: 11 Gaps: 1

 US-09-270-437D-7 (1-1946) x 088477 (1-577)

 QY 108 AGGAGCCGGAATTCGAATATCCGAATATTCACCCAGCCTCCGATGGGAAGTACTGGAC 167
 Db |||||||
 79 ArgSerArgLysIleGlnIleArgAsnIleProGlnLeuArgTrpGluValLeuAsp 98

QY 168 AGCTGTGGTCTCAGTATGGTACAGTAGAGAACTGTGTAGCAAGTGAACACACGAGAGTGA 227
 Db |||||||
 99 SerLeuLeuAlaGlnTyrGlyThrValGluAsnCysGluGlnValAsnThrGluSerGlu 118
 QY 228 ACGCAGTGTGAATGTCCACCTATTCCAAACGGGAGCAGACACGAGCAAGCCATCATGAAG 287
 Db |||||||
 119 ThrAlaValValAsnValThrTyrSerAsnArgGluGlnThrArgGlnAlaIleMetIys 138
 QY 288 CTGAATGGCCACCAAGTGTGGAGAACCATCCCTGGAAGTCTCTCTACATCCCCGATGAGCAG 347
 Db |||||||
 139 LeuAsnGlyHisGlnLeuGluAsnHisAlaLeuIysValSerTyrIleProAspGluGln 158
 QY 348 ATAGCAGAGGACCTGACAATGGGGCCCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGC 407
 Db |||||||
 159 IleThrGlnGlyProGluAsnGlyArgArgGlyGlyPheGlySerArgGlyGlnProArg 178
 QY 408 CAGGGCTCACCTGTGGCAGCGGGGCCCCAGCAAGCAGCAGCAAGTGGACATCCCCCTT 467
 Db |||||||
 179 GlnGlySerProValAlaAlaGlyAlaProAlaIysGlnGlnProValAspIleProLeu 198
 QY 468 CGGCTCCTGTGCCACCCAGTATGTGGTGCCATTATTGGCAAGAGGGGGCCACCATC 527
 Db |||||||
 199 ArgLeuLeuValProThrGlnTyrValGlyAlaIleIleGlyLysGluGlyAlaThrIle 218
 QY 528 CGCAACATCAAAAACACACCCAGTCCCAAGATAGACGTGCATAGAGAGAGACCCAGGT 587
 Db |||||||
 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGly 238
 QY 588 GCAGCTGAAAAAGCCATCAGTGTGCACCTCCACCCCTGAGGGCTGTCTCTCCGCTGTGAAG 647
 Db |||||||
 239 AlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCysSerSerAlaCysLys 258
 QY 648 ATGATCTTGGAGATATCATTAAGAGGCTAAGGACACAAACACCGCTGACGAGGTTCCT 707
 Db |||||||
 259 MetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThrAlaAspGluValPro 278
 QY 708 CTGAAGATCTTGGCCCAATAAATCTTGTAGGGCTCTCATTTGGCAAGAGAGCGGAC 767
 Db |||||||
 279 LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn 298
 QY 768 CTGAAGAGGTAGACCAAGATACCGAGACAAAATCACCATCTCTCTGTTGCAAGACCTT 827
 Db |||||||
 299 LeuLysLysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeu 318
 QY 828 ACCCTTTACAACTGAGAGGACCATCATCTGTGAAGGGGGCCATCGAGAATTGTTGCAAG 887
 Db |||||||
 319 ThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIleGluAsnCysArg 338
 QY 888 GCCGAGCAGGAATAATGAAGAAAGTTCGGAGGGCTATGAGAATGATGTGCTGCCATG 947
 Db |||||||
 339 AlaGluGlnGluIleMetLysLysValArgGluAlaTyrGluAsnAspValAlaAlaMet 358
 QY 948 AGC-----TCTCACCCTGATCCCTGGCTGAAACCTGGCTGTGTAGGTCTTTTCCAGCT 1001
 Db |||||||
 359 SerLeuGlnSerHisIleLeuProGlyLeuAsnLeuAlaAlaValGlyLeuPheProAla 378
 QY 1002 TCATCCAGGAGTCCCGCGCTCCAGCGCTTACTGGGGCTGTCTCCCTATAGTCTCC 1061
 Db |||||||
 379 SerSerSerAlaValProProProSerSerValThrGlyAlaAlaProProSerSer 398
 QY 1062 TTTATGCAAGTCCCGAGCAGAGATGTGCAGGTGTTTATCCCGCCCGCAGGCTGGGC 1121
 Db |||||||
 399 PheMetGlnAlaProGluGlnGluMetValGlnValPheIleProAlaGlnAlaValGly 418
 QY 1122 GCCATCATCGCAAGAGGGCAGCACATCAACAGCTCTCCCGGTTTGGCCAGCGCTCC 1181
 Db |||||||
 419 AlaIleIleGlyLysLysGlyGlnHisIleLysGlnLeuSerArgPheAlaSerAlaSer 438
 QY 1182 ATCAAGATGCAACCCGAAACACCTGACTCCAAAGTTCGTATGTTATCATCTACATGGA 1241
 Db |||||||
 439 IleLysIleAlaProProGluThrProAspSerLysValArgMetValValIleThrGly 458

DR Pfam; PF00076; rtm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS0084; KH TYPE 1; 4.
DR PROSITE; PS0102; RRM; 2.
SQ SEQUENCE 577 AA; 63436 MW; 0647676128PBDLEE CRC64;

Alignment Scores:

Pred. No.: 9.59e-177 Length: 577
Score: 2504.00 Matches: 495
Percent Similarity: 99.40% Conservative: 1
Best Local Similarity: 99.20% Mismatches: 1
Query Match: 70.91% Indels: 2
DB: 11 Gaps: 1

US-09-270-437D-7 (1-1946) x Q8CGX0 (1-577)

QY 108 AGGAGCCGGAATTCMAATCCGAATATTCACCCAGCTCCGATGGGAAGTACTGGAC 167
DB 79 ArgSerArgLysIleGlnIleArgAsnIleProGlnLeuArgTrpGluValLeuAsp 98
QY 168 AGCTGCTGGCTCAGTATGATGAGAGAACTGTGAGCAAGTGAACACCGAGAGTGAG 227
DB 99 SerLeuLeuAlaGlnTyrGlyThrValGluAsnGlnValAsnThrGluSerGlu 118
QY 228 AGCGAGTGGTCAATGTCACCTATTCCACCGGGAGCAGACCGACCCATCATGAG 287
DB 119 ThrAlaValValAsnValThrTyrSerAsnArgGluGlnThrArgGlnAlaIleMetLys 138
QY 288 CTGAATGGCCAGCTGGAGAACCATGCCCCCTGGAAGTCTCTACATCCCCGATGAGCAG 347
DB 139 LeuAsnGlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyrIleProAspGluGln 158
QY 348 ATAGCACAGGGACCTGAGAAATGGCGCCGAGGGGGCTTTGGCTCTCGGGGTGAGCCCGC 407
DB 159 IleAlaGlnGlyProGluAsnGlyArgGlyGlyPheGlySerArgGlyGlnProArg 178
QY 408 CAGGGCTCAGCTGTGGCAGCGGGGGCCCGAGCCAGCAGCAGCAGTGGACATCCCTTT 467
DB 179 GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnValAspIleProLeu 198
QY 468 CGGCTCTCGTGGCCACCCAGCTATGTTGGTGGCTATTATTGGCAAGAGGGGGCCACCATC 527
DB 199 ArgLeuLeuValProThrGlnTyrValGlyAlaIleIleGlyLysGluAlaThrIle 218
QY 528 CGCAACATCACAAACAGACCCAGTCCCAAGATAGAGTGCATGAGAGAGAACCGAGT 587
DB 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGly 238
QY 588 GCAGCTGAAAGGCCATCAGTGTGCACTCCACCCCTGAGGGTGTCTCTCCGCTTGAAG 647
DB 239 AlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCysSerSerAlaCysLys 258
QY 648 ATGATCTTGGAGATTATGCATAAAGGCTAAGGACACCAAAACCGCTGACGAGTTCCC 707
DB 259 MetIleLeuGluIleMethHisLysGluAlaLysAspThrLysThrAlaAspGluValPro 278
QY 708 CTGAAGATCTGGCCCATATTAATCTTTGTAGGGCTCTCATTTGGCAAGGAAGACCGAAC 767
DB 279 LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn 298
QY 768 CTGAAGAAGGTAGACAAGATACCGAGACAAATAATCACCATCTCTCTGTTGCAAGACCTT 827
DB 299 LeuLysLysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeu 318
QY 828 ACCCTTTACAAACCTGAGAGGACCATCATCTGTGAAGGGGGCCATCGAGAAATTTTCAGG 887
DB 319 ThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIleGluAsnCysArg 338
QY 888 GCCGAGCAGAAATATGAGAAGTTCCGAGGCCCTATGAGAAATGATGTGGCTGCCATG 947
DB 339 AlaGluGlnIleMethLysLysValArgGluAlaTyrGluAsnAspValAlaAlaMet 358

QY 948 AGC-----TCTCACCTGATCCCTGGCCTGAACCTGGCTGCTAGAGTCTTTTCCAGCT 1001
DB 359 SerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaValGlyLeuPheProAla 378
QY 1002 TCATCCAGCGCAGTCCCGCCCTCCACGACGCGTTACTGGGGTGTCTCCCTATAGCTCC 1061
DB 379 SerSerSerAlaValProProProSerValThrGlyAlaAlaProTyrGlySer 398
QY 1062 TTTATGAGGCTCCCGAGCAGAGATGTGTGAGGTGTTTATCTCCCGCCAGCAGTGGGC 1121
DB 399 PheMetGlnAlaProGluGlnGluMetValGlnValPheIleProAlaGlnAlaValGly 418
QY 1122 GCCATCATCGCAAGAGGGCAGCACATCAAAAGCTCTCCCGTTGGCCAGCGCTCC 1181
DB 419 AlaIleIleGlyLysGlyGlnHisIleLysGlnLeuSerArgPheAlaSerAlaSer 438
QY 1182 ATCAAGATTGACCAACCCGAAACACCTGACTCCAAAGTTTCGTATGTTATCATCTGGA 1241
DB 439 IleLysIleAlaProProGluThrProAspSerLysValArgMetValIleThrGly 458
QY 1242 CGCCAGAGGCCCAATTCAGGCTCAGGAGAGAAATCTATGGCAAACTCAAGGAGGAC 1301
DB 459 ProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysLeuLysGluAsn 478
QY 1302 TTCTTTGGTCCCAAGGAGAGTGAAGCTGGAGACCCACATACGTGTGCCATCAGCA 1361
DB 479 PhePheGlyProLysGluGluValLysLeuGlnThrHisIleArgValProAlaSerAla 498
QY 1362 GCTGCGCGGCTCATTTGGCAAGGTGGAACCGGTGAACGAGTTGCAGAAATTTGACGGA 1421
DB 499 AlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThrAla 518
QY 1422 GCTGAGGTGCTAGTACCAAGACAGACCCCTGTATGAGACGACGAGTCTCTGTA 1481
DB 519 AlaGluValValProArgAspGlnThrProAspGluAsnAspGlnValIleValLys 538
QY 1482 ATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCTGGCC 1541
DB 539 IleIleGlyHisPheTyrAlaSerGlnMetAlaGlnArgLysIleArgAspIleLeuAla 558
QY 1542 CAGGTTAAGCAGCAGCATCAGAGGAGGACAGATAACCGGCCCGCCAGGAGGAG 1598
DB 559 GlnValLysGlnGlnHisGlnLysGlyGlnSerAsnGlnAlaGlnAlaArgArgLys 577

RESULT 2

Q9NZ18 PRELIMINARY; PRT; 577 AA.
AC Q9NZ18; 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE mRNA-binding protein CRBP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ioannidis P., Trangas T., Dimitriadis E., Samiotaki M.,
RA Panoutsakopoulos G., Kyriazoglou I., Voutzoulis S., Tsiapalis C.M.,
RA Kittas C., Agnantis N., Pandis N.;
RT "Ectopic expression of a KH-domain containing protein, highly
RT homologous to both human IMP-1 and mouse CRD-BP, in benign and
RT malignant mesenchymal tumors";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AF198254; AAF37203.1; -.
DR HSSP; PI1940; 1CVJ.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH_dom.
DR InterPro; IPR004088; KH_type_1.
DR InterPro; IPR00504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 23, 2004, 10:55:00 ; Search time 93.0388 Seconds
(without alignments)
13198.757 Million cell updates/sec

Title: US-09-270-437D-7

Perfect score: 2531
Sequence: 1 gcctagcgaggggctggg.....attcttcagggttttaaaa 1946

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 2034082

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_SPOOL_p/US09270437/runat_23072004_095739_27425/app_query.fasta_1.5582
-DB=SPTREMBL_25 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09270437 @CGN 1 1 287 @runat_23072004_095739_27425 -NCPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SPTREMBL 25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2504	70.9	577	11	Q8CGX0 Q8CGX0 rattus norv

2	2499	70.8	577	4	Q9NZI8
3	2491	70.5	577	11	Q88477
4	2484	70.3	577	11	Q8BRH1
5	2484	70.3	577	11	Q80US9
6	2389.5	67.7	576	13	Q42254
7	1975.5	55.9	594	13	Q73932
8	1972	55.8	593	13	Q57526
9	1873	53.0	579	4	Q00425
10	1864	52.8	579	11	Q9CPN8
11	1857	52.6	579	11	Q8C2J9
12	1830.5	51.8	582	13	Q9PW80
13	1697	48.1	556	4	Q9Y6M1
14	1299	36.8	545	11	Q7TQF9
15	863.5	24.5	566	5	Q9VZ69
16	863.5	24.5	573	5	Q8IR99
17	859	24.3	580	5	Q8IGK4
18	466.5	13.2	828	5	Q21605
19	265	7.5	621	10	Q9C553
20	257.5	7.3	169	11	Q7TP50
21	256.5	7.3	398	3	Q74919
22	251	7.1	454	10	Q84NA6
23	251	7.1	479	10	Q9XI71
24	248	7.0	557	5	Q23487
25	246.5	7.0	510	5	Q7Z145
26	242	6.9	318	13	Q7ZVK5
27	241.5	6.8	672	13	Q8AX85
28	240.5	6.8	680	5	P91393
29	238	6.7	542	10	Q8S7G1
30	237.5	6.7	762	10	Q8LNT9
31	236.5	6.7	641	5	Q9BLA0
32	236.5	6.7	653	13	Q7ZXS1
33	235	6.7	577	10	Q9SR13
34	234.5	6.6	568	10	Q9LXF5
35	233.5	6.6	313	4	Q96EP6
36	232.5	6.6	632	10	Q82762
37	232	6.6	589	5	Q17935
38	232	6.6	611	5	Q17936
39	231	6.5	640	10	Q9ASX3
40	231	6.5	644	10	Q9PNK3
41	230	6.5	774	10	Q9L128
42	228	6.5	644	10	Q8LDV1
43	223.5	6.3	610	13	Q8AX84
44	222.5	6.3	364	5	Q95SZ9
45	220.5	6.2	353	13	Q9W702

ALIGNMENTS

RESULT 1

ID	Q8CGX0	PRELIMINARY;	PRT;	577 AA.
AC	Q8CGX0;			
DT	01-MAR-2003 (TrEMBLrel. 23, Created)			
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)			
DE	B-actin zipcode binding protein 1.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCHI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Sprague-Dawley;			
RT	Eom T., Singer R.H., Bassell G.J.;			
RT	"Molecular interactions between rzbp1 and b-actin zipcode required for			
RT	transport of mRNA and stimulation of spine growth."			
RL	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF541940; AA016210.1;			
DR	GO; GO:0003676; F:nucleic acid binding; IEA.			
DR	InterPro; IPR004087; KH dom.			
DR	InterPro; IPR004088; KH type 1.			
DR	InterPro; IPR000504; RNA_rec_mot.			
DR	Pfam; PF00013; KH; 4.			

This Page Blank (uspto)

```
Db 418 nGluileArgLysArgThrAsnAlaAsnIleCysIleSerLysGlyLysLysAspLe 438
QY 827 TACCCTTTACACCTGAGAGGACATCAGTGTGAAGGGGCCCATCGAGAATGTGCGAG 886
Db 438 u-----ValGluValSerGlyGluValSerSerValArgAs 450
QY 887 GGCGAGCAGGAATAATGAAGAAGTTCGGAGGCCCTATGAGAATGATGTGGTGCCAT 946
Db 450 pAlaLeuIleGlnIleValLeuArgLeuArgGluAspValLeuGlyAspLysAspSerVa 470
QY 947 GAGCTCTCACCTGATCCCT----- 965
Db 470 lAlaThrArgLysProProAlaArgThrAspAsnCysSerPheLeuSerGlySerSerAs 490
QY 966 ----GGCCTGAACCTGGCTGTGTAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCG-- 1019
Db 490 nAlaGlyTyrThrLeuProSerPheMetSerSerMetAlaSerThrSerGlyPheHisGl 510
QY 1020 -----CGCCTCCAGCAGCGTTACTGGGGCTGCTCCCTATAGCTCTTTAT 1066
Db 510 yTyrGlySerPheProAlaGlyAspAsnValLeuGlySerThrGlyProTyrSerTyrGl 530
QY 1067 GCAGGCTCCCGAGCAGGAGATGGTGCGAGTGTGTTATCCCGCCCGCAGTGGCGCCAT 1126
Db 530 yArgLeuProSerSerSerAlaLeuGluLeuLeuProAlaHisAlaMetSerLysVa 550
QY 1127 CATCGCAAGAAGGGGCGACACATCAACAGCTCTCCCGGTTTCCAGCGCCTCCATCAA 1186
Db 550 lMetGlyLysGlyGlyGlyAsnLeuGluAsnIleArgArgIleSerGlyAlaMetIleGl 570
QY 1187 GATTGCACCCCGAACA 1205
Db 570 uIleSerAlaSerLysThr 576
```

Search completed: July 23, 2004, 11:06:27
Job time : 41.1191 secs

RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA Neumann S., Argirious A., Vitale D., Liguori R., Piravandi E.,
RA Maesenet O., Quigley F., Clabaud G., Muendlein A., Felber R.,
RA Schnabl S., Hiller K., Berger C., Monfort A., Casacuberta E.,
RA Chefdor F., Cooke R., Weber N., Vandenberg M., Barges M., Terol J., Torres A.,
RA Gibbons T., Weber N., Vandenberg M., Barges M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francs P., Bieleke C.,
RA Prishan D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threiden J.,
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.W.,
RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshua C.,
RA Antonoli B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA Chen E., Marra M.A., Martienssen R., McCombie W.R.; the
RA "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RA thaliana.";
RA Nature 402:769-777(1999).
RA [2]
RA SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RA STRAIN=cv. Columbia;
RA MEDLINE=22954850; PubMed=14593172;
RA Yamada K., Lim J.M., Dale J.M., Chen H., Shinn P., Palm C.J.,
RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.F., Carninci P.,
RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA Satou M., Tamez R., Vayssier M., Wallender E.K., Wong C., Yamamura Y.,
RA Yuan S., Shinohara K., Davis R.W., Theologis A., Ecker J.R.;
RA "Empirical analysis of transcriptional activity in the Arabidopsis
RA genome.";
RA Science 302:842-846(2003).
RA CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
RA CC -!- ALTERNATIVE PRODUCTS:
RA CC Event=Alternative splicing; Named isoforms=2;
RA CC Name=1;
RA CC IsoId=P58223-1; Sequences=Displayed;
RA CC Note=No experimental confirmation available;
RA CC Name=2;
RA CC IsoId=P58223-2; Sequences=VSP_008899, VSP_008900;
RA CC Note=May be due to a competing acceptor site. No experimental
RA CC confirmation available;
RA CC -!- SIMILARITY: Contains 5 KH domains.
RA CC -!- CAUTION: Ref.1 sequences differ from that shown due to erroneous
RA CC gene model prediction. At4g18370 and At4g18375 were originally
RA CC fused into a single gene.
RA CC -----
RA CC This SWISS-PROT entry is copyright. It is produced through a collaboration
RA CC between the Swiss Institute of Bioinformatics and the EMBL outstation
RA CC at the European Bioinformatics Institute. There are no restrictions on its
RA CC use by non-profit institutions as long as its content is in no way
RA CC modified and this statement is not removed. Usage by and for commercial
RA CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
RA CC or send an email to license@isb-sib.ch).
RA CC -----
RA CC EMBL; AL021710; CAA16717.1; ALT_SEQ.
RA CC EMBL; AL161548; CAB78839.1; ALT_SEQ.
RA CC EMBL; AY133701; AAM91635.1; -.
RA CC EMBL; BT001108; AAN64172.1; -.
RA CC InterPro; IPR004087; KH_dom.

DR InterPro; IPR004088; KH_type_1.
DR Pfam; PF00013; KH; 5.
DR SMART; SM00322; KH; 5.
DR PROSITE; PS00084; KH TYPE 1; 5.
KW Hypothetical protein; Nuclear protein; RNA-binding; Repeat;
KW Alternative splicing 99
FT DOMAIN 35 KH 1.
FT DOMAIN 138 210 KH 2.
FT DOMAIN 311 380 KH 3.
FT DOMAIN 394 455 KH 4.
FT DOMAIN 535 599 KH 5.
FT VARSPLIC 532 532 L -> F (in isoform 2).
FT VARSPLIC 533 606 Missing (in isoform 2).
FT VARSPLIC 533 606 /FTID=VSP_008899.
FT SEQUENCE 606 AA; 65760 MW; 61F135BBB8647C0C CRC64;
SQ
Alignment Scores: 3.98e-06 Length: 606
Pred. No.: 216.50 Matches: 89
Score: 43.67% Conservative: 80
Percent Similarity: 23.00% Mismatches: 164
Best Local Similarity: 6.13% Indels: 54
Query Match: 1 Gaps: 10
DB:
US-09-270-437d-7 (1-1946) x Y475_ARATH (1-606)
QY 125 AATCGAAATATTCACCCAGCTCCGATGGGAAGTACTGACAGCGCTGGTGGCTCAGTA 184
DB 218 AsnProArgGluAsnIleProLeuAspSerThrSerGlnAspValProAlaAlaSerVal 237
QY 185 TGTACACTAGAGAACTGTGACAGTCAACACCGAGAGTGGAGCGG-----AGTGGT 238
DB 238 IleValProSerAspLeuSerAsnSerValTyrProGln---ThrGlyPheTyrSerAsn 256
QY 239 GAATGTCACCTATTCCACCGGAGCAGACAGCCAGCCATCATGAAGCTGAATGGCCA 298
DB 257 GlnAspHisIleLeuGlnGlnGlyAlaGlyValProSerTyrPheAsnAla----- 273
QY 299 CCAGTTGGAGAACCATGCCCTGGAAGGTCCTACATCCCCGATGAGCAGATAGCAGCGG 358
DB 274 -----LeuSerValSerAspPheGlnGly 281
QY 359 ACCTGAGATGGCGCCGCGGGGGTGGTCTCGGGGTGACGCCCGCGAGGCTCAC 418
DB 282 TyrAlaGluThrAlaAlaAsnProValProValPheAlaSerSerLeuPro---ValThr 300
QY 419 TGTGCGAGCGGGGCGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 478
DB 301 HisGlyPheGlyGlySerSer-ArgSerGlu-----GluLeuValPheLeuValLeuCy 318
QY 479 GCCACCCAGTATGTGGTGCCATTTGGCAAGAGGGGGGCCACCATCCGCAACATCAC 538
DB 318 sProLeuCyAsnIleMetArgValIleGlyGlySerThrThrIleLeuArgIle 338
QY 539 AAAACAGACCCAGTCCAGATAGAGCTGATAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 598
DB 338 gGluAlaSerGlySerCysIleGluValAsnAspSerArgThrLysCysGlyAspAspG 358
QY 599 AGCCATCAGTGTGCACTCCACCCCTGCTCTCCCTGCTGTGTAGATGATCTTGA 658
DB 358 uCysValIleIleValThrAlaThrGluSerProAspMetLysSerMetAlaValG 378
QY 659 G-----ATTATGCAATAAGAGCTTAAGACACCAACACCGCTGACGAGTTCCTCGAA 712
DB 378 uAlaValLeuLeuGlnGlnTyrIleAsnAspGluAspAlaGluAsnValLysMetG 398
QY 713 GATCTGGGCCATAATACTTTGTAGGGGCTCTCATTTGGCAGGAGGAGGAGGAGGAGGAG 772
DB 398 nLeuLeuValSerSerLysValIleGlyCysValIleGlyLysSerGlySerValIleAs 418
QY 773 GAAGGTAGAGCAAGATACCGAGACACAAAATCACCATCTCC-----TCGTTGCAAGACCT 826
DB 773 GAAGGTAGAGCAAGATACCGAGACACAAAATCACCATCTCC-----TCGTTGCAAGACCT 826


```
Db 443 AspHisLysPheHisArgHisLeuIleGlyLysSerGlyAlaAsnIleAsnArgIleLys 462
QY 339 GATGAG-----CAGATACACAGGACCTGAGAATGGCGCGGAGGGGCTTTGGC 389
Db 463 AspGlnTyrLysValSerValArgIleProPheAspSerGluLysSerAsnLeuIleArg 482
QY 390 TCTCGGGTACGCCCGCCAGGCTCACCTGTGGCAGCGGGGGCCCGCAGCCAGCAGCAG 449
Db 483 IleGluGlyAspPro---GlnGlyVal-----GlnGlnAlaLysArgGlu 496
QY 450 CAAGTGGACATCCCTCTCGG-----CTCTGGTGGCCACC 485
Db 497 LeuLeuGluLeuAlaSerArgMetGluAsnGluArgThrLysAspLeuIleIleGluGln 516
QY 486 CAGTATGTGGGTGCCATTATTGGCAAGGAGGGGCCACCATCCGCAACATCACAAACAG 545
Db 517 ArgPheHisArgThrIleIleGlyGlnLysGlyGluArgIleArgGluIleArgAspLys 536
QY 546 ACCCAGTCCAAGATA-----GACGTGCATAGGAAGGAGAACGCGAGTGCAGT 593
Db 537 PheProGluValIleIleAsnPheProAspProAlaGlnLysSerAsp----- 552
QY 594 GAAAAAGCCATCAGTGTGCACCTCCACCCCTGAGGGCTGCTCCTCGCTGTGAAGATGTC 653
Db 553 -----IleValGlnLeuArgGlyProLysAsnGluValGluLysCysThrLysTyr 569
QY 654 TTGGAGATTATCATAGAGGCTAAGGACACCAAAACGGCTGAGAGGTTCCCTGTAAG 713
Db 570 MetGlnLysMetValAlaAspLeuValGluAsnSerTyrSerIleSerValProIlePhe 589
QY 714 ATCCTGGCCCAATAATACCTTTGTAGGGCTCTCATTTGGCAAGGAGCGGACCTGAG 773
Db 590 LysGlnPheHisLysAsn-----IleIleGlyLysGlyAlaAsnIleLys 605
QY 774 AAGGTAGACAGATACCGAGACAAATAATCACCATCTCTCTGTCGTAAGACCTTACCTT 833
Db 606 LysIleArgGluGluSerAsnThrLysIle-----AspLeuProAla 619
QY 834 TACACACCT---GAGAGGACCATCATCTGTAGAGGGGCCCATCAGAAATTTGTGAGGGCC 890
Db 620 GluAsnSerAsnSerGluThrIleIleIleThrGlyLysArgAlaAsnCysGluAlaAla 639
QY 891 GAGCAGGAATAATGAGAAAGTTCGGGAGGCTATGAGAAATGATGTGCTGCCATGAGC 950
Db 640 ArgSerArgIleLeu-----SerIleGlnLysAspLeuAla----- 651
QY 951 TCTCACTGATCCTCGCTGCGCTGAACCTGCTGCTGTAGTCTTTTCCAGCTTCATCCAGC 1010
Db 652 -----AsnIleAlaGlu----- 655
QY 1011 GCAGTCCCGCGCTCCAGACGCGTTACTGGGGCTGCTCCCTATAGTCCCTTTATGAG 1070
Db 655 ----- 655
QY 1071 GCTCCCGAGCAGGATGTTGTCAGGTTTATCCCGCCCGCAGCAGTGGCGCATCATC 1130
Db 656 -----ValGluValSerIleProAlaLysLeuHisAsnSerLeuIle 669
QY 1131 GGCAAGAGGGGCGAGCACATCAACAGCTCTCCCGGTTTGGCCAGCGCTCCCATCAGATT 1190
Db 670 GlyThrLysGlyArgLeuIleArgSerIleMetGluGluCysGlyGlyValHisIleHis 689
QY 1191 GCACCAACCCCAACACCTGACTCAAGTTCGATGTTATCATCTATGACCGCCAGAG 1250
Db 690 PheProValGluGlySerGlySerAspThr-----ValValIleArgGlyProSerSer 707
QY 1251 GCCCAATTCAAGCTCAGGAGAAATCTATGCGAAACTCAAGCAGGAGACTTCTTGGT 1310
Db 708 AspValGluLysAlaLysLysGlnLeuLeu---HisLeuAlaGlu----- 722
QY 1311 CCCAAGGAGGAAGTGAAGCTGGAGACCCACATACATCGTGTGCCAGCATCAGCAGTGGCCGG 1370
|||||
```

```
Db 723 ---LysGlnThrLysSerPheThrValAspIleArgAlaLysProGluTyrHisLysPhe 741
QY 1371 GTCATTGGCAAGGTGGAAAAACCGTGAAACGAGTTGCAGATTTCACGCGCAGCTCAGGTG 1430
Db 742 LeuIleGlyLysGlyGlyLysIleArgLysValArgAspSerThrGlyAlaArgVal 761
QY 1431 GTAGTACCACGAGAGACGACCCCTGATGAGAACGACGAGGTCTCGTGAATAATCATCGGA 1490
Db 762 IlePhePro-----AlaAlaGluAspLysAspGlnAspLeuIleThrIleIleGly 778
QY 1491 CATTTCTATGTCAGTCAGATGGCTCAACGGAAGATCCGAGACATCTCGGCCAGGTTAAG 1550
Db 779 LysGluAspAlaValArgGluAlaGlnLysGluLeuGluAlaLeuIle-GlnAsnLeuAs 798
QY 1551 CAGCAGCATCAGAGGGACAGAGTAACGAGGCCCGCCAGGCAC 1590
Db 798 pAsnValValGluAspSerMetLeuValAspProLysHis 811
RESULT 14
PCB4_HUMAN
ID PCB4_HUMAN STANDARD; PRT; 403 AA.
AC P57723;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Poly(rc)-binding protein 4 (Alpha-CP4).
GN PCB4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
SEQUENCE FROM N.A.
RP MEDLINE=20396135; PubMed=10936052;
RA Makeyev A.V., Liehaber S.A.;
RT "Identification of two novel mammalian genes establishes a subfamily
of KH-domain RNA-binding proteins.";
RL Genomics 67:301-316(2000).
[2]
SEQUENCE FROM N.A.
RP Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saio K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
RP TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.C., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.W., Krzywinski M.I., Schmitz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
```

Db 238 LeuAsnGlnValAlaArgGlnGlnSerHisPheAlaMetMetHisGlyGlyThrGlyPhe 257
QY 1029 AGCAGCGTACTGGGGTGTCTCC-----TATAGTCTCTTTATGACGCTCCC 1076
Db 258 AlaglyIleAspSerSerProGluValIleGlyTyrTrpAlaSerLeuAspAlaSer 277
QY 1077 GAGCAGAGATGGTGCAGGTGTATATCCCGCCAGCAGTGGGGCCATCATCGGCAAG 1136
Db 278 ThrGlnThrThrHisGluLeuThrIleProAsnAsnLeuIleGlyCysIleIleGlyArg 297
QY 1137 AAGGGGAGCAGACACAGCTCTCCGGTTCGCCAGCTCCATCAAGATGACCA 1196
Db 298 GlnGlyAlaAsnIleAsnGlnIleArgGlnMetSerGlyAlaGlnIleIleValAlaAsn 317
QY 1197 CCCGAAACCTGACTCAAAAGTTCGTATGTTATCATCTGACCGCCAGCGCCCAA 1256
Db 318 Pro---ValGluGlySerSerGlyArgGlnValThrIleThrGlySerAlaAlaSerIle 336
QY 1257 TTCAGGCTCAGGAGAACTATGCGAACTCAAGGAGGAGAACTCTTTGGT 1310
Db 337 SerLeuAlaGlnIleThrLeuIleAsnAlaArgLeuSerSerGlyGlyMetGly 354

RESULT 13

VGLN_HUMAN

ID VGLN_HUMAN STANDARD; PRT; 1268 AA.
AC Q00341; Q9UCY3;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Vlgilin (High density lipoprotein-binding protein) (HDL-binding
DE protein).
GN HDLBP OR HBP OR VGL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92291094; PubMed=1318310;
RA McKnight G.L., Reasoner J., Gilbert T., Sundquist K.O., Hokland B.,
RA McKernan P.A., Champagne J., Johnson C.J., Bailey M.C., Holly R.,
RA O'Hara P.J., Oram J.F.,
RT "Cloning and expression of a cellular high density lipoprotein-binding
RT protein that is up-regulated by cholesterol loading of cells.";
RL J. Biol. Chem. 267:12131-12141(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SUBCELLULAR LOCATION.
RX MEDLINE=96184515; PubMed=8605996;
RA Kugler S., Grunweller A., Probst C., Klinger M., Muller P.K.,

RA Kruse C.;
RT "Vlgilin contains a functional nuclear localisation sequence and is
RT present in both the cytoplasm and the nucleus.";
RL FEBS Lett. 382:330-334(1996).
CC -!- FUNCTION: APPEARS TO PLAY A ROLE IN CELL STEROL METABOLISM.
CC IT MAY FUNCTION TO PROTECT CELLS FROM OVER-ACCUMULATION OF
CC CHOLESTEROL.
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
CC -!- SIMILARITY: Contains 14 KH domains.
CC -!- SIMILARITY: STRONG, TO C.ELEGANS C08H9.2.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M64098; AAA35962.1; -.
DR EMBL; BC001179; AAH01179.1; -.
DR PIR; A44125; A44125.
DR PDB; 1VIG; 03-APR-96.
DR PDB; 1VIR; 03-APR-96.
DR Genew; HGNC:4857; HDLBP.
DR MIM; 142695; -.
DR GO; GO:0005886; C:plasma membrane; TAS.
DR GO; GO:0008289; F:lipid binding; TAS.
DR GO; GO:0008203; P:cholesterol metabolism; TAS.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH_type_1.
DR Pfam; PF00013; KH; 14.
DR SMART; SM00322; KH; 12.
DR PROSITE; PS00084; KH type 1; 14.
KW HDL; Lipid transport. Cholesterol metabolism; RNA-binding; Repeat;
KW Phosphorylation; Nuclear protein; 3D-structure.
FT DOMAIN 158 229 KH 1.
FT DOMAIN 230 302 KH 2.
FT DOMAIN 303 371 KH 3.
FT DOMAIN 372 442 KH 4.
FT DOMAIN 443 514 KH 5.
FT DOMAIN 515 588 KH 6.
FT DOMAIN 589 660 KH 7.
FT DOMAIN 661 734 KH 8.
FT DOMAIN 735 807 KH 9.
FT DOMAIN 808 880 KH 10.
FT DOMAIN 881 979 KH 11.
FT DOMAIN 980 1059 KH 12.
FT DOMAIN 1060 1134 KH 13.
FT DOMAIN 1135 1209 KH 14.
FT MOD_RES 295 295 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 296 296 PHOSPHORYLATION (POTENTIAL).
SQ SEQUENCE 1268 AA; 141439 MW; 374ECB83D13A7431 CRC64;
Alignment Scores:
Pred. No.: 3,85e-06 Length: 1268
Score: 218.00 Matches: 109
Percent Similarity: 38.46% Conservative: 81
Best Local Similarity: 22.06% Mismatches: 191
Query Match: 6.17% Indels: 113
DB: 1 Gaps: 19
US-09-270-437d-7 (1-1946) x VGLN_HUMAN (1-1268)
QY 195 GAGAACTGTGAGCAAGTGAACACGAGAGTGAGAGCGAGTGTGAATGTCACCTATTC 254
Db 403 GluGlyGluAspIleThrLeuGluGlyProThrGluAspValAsnValAlaGlnGlu 422
QY 255 AACCGGAGCAGC-----ACCAGGCAAGCCATCATGAAGTGAAT--- 293
Db 423 GlnIleGluGlyMetValIleAsnArgMetAspTyrValGluIleAsnIle 442
QY 294 GGCACCAAGTTGGAGAACCATGCCCTG-----AAGTCTCTTACATCCCC 338

RA Kavsian V.M.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RX X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 423-495.
RF MEDLINE=99148126; PubMed=10368286;
RA Lewis H.A., Chen H., Edo C., Buckanovich R.J., Yang Y.Y.-L.,
RA Musumuru K., Zhong R., Darnell R.B., Burley S.K.;
RT "Crystal structures of Nova-1 and Nova-2 K-homology RNA-binding
domains";
RL Structure 7:191-203 (1999).
CC -!- FUNCTION: May regulate RNA splicing or metabolism in a specific
subset of developing neurons.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=P51513-1; Sequence=Displayed;
CC Name=2; Synonyms=Tumor;
CC IsoId=P51513-2; Sequence=VSP_002841, VSP_002843, VSP_002844;
CC Name=3;
CC IsoId=P51513-3; Sequence=VSP_002842;
CC -!- TISSUE SPECIFICITY: Brain.
CC -!- DISEASE: Autoantigen in the paraneoplastic opsoclonus myoclonus
ataxia (POMA), a paraneoplastic neurological syndrome/disorder
(PNS/D) associated with breast cancer, fallopian cancer, and SCLCa
and characterized primarily by loss of inhibitory control of motor
neurons in the spinal cord and brainstem. Recognized by the IgG
antibody ANNA-2 (also called anti-Ri).
CC -!- SIMILARITY: Contains 3 KH domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U04840; AAA16022.1; -;
DR EMBL; Z70771; CAA94810.1; -;
DR PIR; I38489; I38489.
DR PDB; 1DT4; 10-MAY-00.
DR Genew; HGNC:7886; NOVAL.
DR GK; P51513; -;
DR MIM; 602157; -;
DR GO; GO:0003723; F:RNA binding; TAS.
DR GO; GO:0007626; P:Locomotor behavior; TAS.
DR GO; GO:0008380; P:RNA splicing; TAS.
DR GO; GO:0007268; P:synaptic transmission; TAS.
DR InterPro; IPR004087; KH_dom.
DR Pfam; PF00013; KH; 3.
DR SMART; SM00322; KH; 3.
DR PROSITE; PS50084; KH_TYPE_1; 3.
KW Antigen; Nuclear protein; RNA-binding; Repeat; Alternative splicing;
KW 3D-structure.
FT DOMAIN 27 43 BIPARTITE NUCLEAR LOCALIZATION SIGNAL
(POTENTIAL).
FT DOMAIN 49 119 KH 1.
FT DOMAIN 174 240 KH 2.
FT DOMAIN 276 412 ALA-RICH.
FT DOMAIN 424 491 KH 3.
FT VARSPLIC 88 90 Missing (in isoform 2).
FT VARSPLIC 153 176 Missing (in isoform 3).
FT VARSPLIC 177 184 Missing (in isoform 2).
FT VARSPLIC 185 510 Missing (in isoform 2).
FT VARSPLIC 185 510 Missing (in isoform 2).
FT VARSPLIC 185 510 Missing (in isoform 2).
SQ SEQUENCE 510 AA; 52055 MW; 7B9BA8B67CAB6984 CRC64;

Alignment Scores:

	Pred. No.:	1.86e-06	Length:	510
Score:	221.50	Matches:	111	
Percent Similarity:	34.89%	Conservative:	68	
Best Local Similarity:	21.64%	Mismatches:	173	
Query Match:	6.27%	Indels:	161	
DB:	1	Gaps:	17	

US-09-270-437d-7 (1-1946) x NOAL_HUMAN (1-510)

Qy	363	GAGATGGCGCGCGAGGGGCTTT	-----GGCTCTCGGGGTGAC	401
Db	9	GlnAsnGlyThrHisThrGlyValProIleAspLeuAspProProAspSerArgLysArg		28
Qy	402	CCCCCGCAGGGCTCACTGTGTGCAGCGGGGCCCAAGCAGCAGCAAGTGGAC	-----	458
Db	29	ProLeuGluAlaProProGluAlaGlySerThrLysArgThrAsnThrGlyGluAspGly		48
Qy	459	---ATCCCCCTCGGCTCTGGTCCGCCACCAGTATGTGGTGCATATTGGCAGAG	515	
Db	49	GlnTyrPheLeuLysValLeuIleProSerTyrAlaAlaGlySerIleGlyGly	68	
Qy	516	GGGGCCCATCCGCAACATCACAAAACAGACCAGTCCCAAGATAGACGTGCATAGGAG	575	
Db	69	GlyGlnThrIleValGlnLeuGlnLysGluThrGlyAlaThrIleLysLeuSerLysLeu	88	
Qy	576	GAGAACGCA	-----GGTGCAGCTGAAAGCCATCAGTGTGCATCTCCACC	620
Db	89	SerLysSerLysAspPheTyrProGlyThrThrGluArgValCysLeuIleGlnGlyThr	108	
Qy	621	CCTGAGGGCTCTCTCGCT	-----TGTAAGATGATCTTGGAGATATGAT	668
Db	109	ValGluAlaLeuAsnAlaValHisGlyPheIleAlaGluLysIleArgGluMetProGln	128	
Qy	669	AAAGAGCTAAG	-----	680
Db	129	AsnValAlaLysThrGluProValSerIleLeuGlnProGlnThrThrValAsnProAsp	148	
Qy	681	-----GACACCAAAACGGTGCAGAGTTCCC	707	
Db	149	ArgIleLysGlnThrLeuProSerSerProThrThrThrLysSerSerProSerAspPro	168	
Qy	708	CTG	-----AAGATCTGCGCCCATATACTTTGTGGGCGT	743
Db	169	MetThrThrSerArgAlaAsnGlnValLysIleIleValProAsnSerThrAlaGlyLeu	188	
Qy	744	CTCATTCGCAAGGAGGACGACCTCAAGAGGTAGCAGCAAGATACCGACACAAATC	803	
Db	189	IleIleGlyLysGlyGlyAlaThrValLysAlaValMetGlnGlnSerGlyAlaTrpVal	208	
Qy	804	ACCATCTCTCTGTTCAGACCTTACCCCTTTACAAACCTGAGAGGACCATCACTGTGAAG	863	
Db	209	GlnLeuSerGlnLysProAsp	---GlyIleAsnLeuGlnGluArgValValThrValSer	227
Qy	864	GGGGCCCATCAGAAATTTGTCAGGGCCGACGAGGAATAATGAAGAAGTTCGGAG	920	
Db	228	GlyGluProGlnGlnAsnArgLysAlaValGluLeuIleIleGlnLysIleGlnGluAsp	247	
Qy	921	-----GCCTATGAGAA	932	
Db	248	ProGlnSerGlySerCysLeuAsnIleSerTyrAlaAsnValThrGlyProValAlaAsn	267	
Qy	933	-----GATGGTGGTCCCATG	947	
Db	268	SerAsnProThrGlySerProTyrAlaAsnThrAlaGluValLeuProThrAlaAla	287	
Qy	948	AGCTCTCACCTGATCCCTGCGCCCTGACCTGGCTGTAGTCTTTTCCCA	998	
Db	288	AlaAlaGlyLeuLeuGlyHisAlaAsnLeuAlaGlyValAlaAlaPheProAlaValLeu	307	
Qy	998	-----	998	
Db	308	SerGlyPheThrGlyAsnAspLeuValAlaIleThrSerAlaLeuAsnThrLeuAlaSer	327	

FT DOMAIN 13 29 BIPARTITE NUCLEAR LOCALIZATION SIGNAL
 FT DOMAIN 35 102 (POTENTIAL).
 FT DOMAIN 157 223 KH 1.
 FT DOMAIN 259 395 KH 2.
 FT DOMAIN 407 474 ALA-RICH.
 SQ SEQUENCE 493 AA; 50302 MW; 7D09E8A55B0A7817 CRC64;

Alignment Scores:
 Pred. No.: 1.84e-06 Length: 493
 Score: 221.50 Matches: 107
 Percent Similarity: 35.6% Conservative: 69
 Best Local Similarity: 21.6% Mismatches: 167
 Query Match: 6.27% Indels: 151
 DB: 1 Gaps: 16

US-09-270-437D-7 (1-1946) x NOA1_MOUSE (1-493)

QY 390 TCTCGGGTCCAGCCCGCCAGGGCTCAGCTGTGGCAGCGGGGGCCCGCCAGCAGCAGCAG 449
 DB 11 SerArgLysArgProLeuGluAlaProGluAlaGlySerThrLysArgThrAsnThr 30
 QY 450 CAAGTGGAC-----ATCCCCCTTCGGCTCCTGGTCCCGCCAGCCAGTATGTGGTGCCATT 503
 DB 31 GlyGluAspGlyGlnTyrPheLeuLysValLeuIleProSerTyrAlaAlaGlySerIle 50
 QY 504 ATTGCAAGAGGGGGCCACCATCGCAGACATCAAAACAGACACCAGTCCAGATAGAC 563
 DB 51 IleGlyLysGlyGlyGlnThrIleValGlnLeuGlnLysGluThrGlyAlaThrIleLys 70
 QY 564 GTGCATAGAGGAGAAC-----CGAGTGCAGCTGAAAGCCCATCAGTGTGCACCTCC 617
 DB 71 LeuSerLysSerLysAspPheTyrProGlyThrGluArgValCysLeuIleGlnGly 90
 QY 618 ACCCTCGAGGGTGTCTCTCCGCT-----TGTAAGATGATCTGGAGATTATG 665
 DB 91 ThrIleGluAlaLeuAsnAlaValHisGlyPheIleAlaGluLysIleArgGluMetPro 110
 QY 666 CATAAGAGGCTAAG-----GACACCAAAACAGCGTGCAGAGGTT 704
 DB 111 GlnAsnValAlaLysThrGluProValSerIleLeuGlnProGlnThrThrValAsnPro 130
 QY 681 -----GACACCAAAACAGCGTGCAGAGGTT 704
 DB 131 AspArgIleLysGlnThrLeuProSerSerProThrThrLysSerSerProSerAsp 150
 QY 705 CCCTG-----AGATCTGGCCCATTAATCACTTTGTAGGG 740
 DB 151 ProMetThrThrSerArgAlaAsnGlnValLysIleIleValProAsnSerThrAlaGly 170
 QY 741 CGTCTATTGGCAAGAGCGGACCTGAGACGAGTACAGCAATACACCGACAAAA 800
 DB 171 LeuIleIleGlyLysGlyGlyAlaThrValLysAlaIleMetGluGlnSerGlyAlaTrp 190
 QY 801 ATCCATCTCTCTGTTGCAAGACTTACCTTCAACCCCTGAGAGGACCACTCATGTG 860
 DB 191 ValGlnLeuSerGlnLysProAsp---GlyIleAsnLeuGlnGluArgValThrVal 209
 QY 861 AAGGGGCCCATCAGAAATTGTCAGCGCGGACGAGAAATATAGAAAGATTCGGGAG 920
 DB 210 SerGlyGluProGluGlnAsnArgLysAlaValGluLeuIleIleGlnLysIleGlnGlu 229
 QY 921 -----GCTATGAGNAT----- 932
 DB 230 AspProGlnSerGlySerCysLeuAsnIleSerTyrAlaAsnValThrGlyProValAla 249
 QY 933 -----GATGTGGTGGC 944
 DB 250 AsnSerAsnProThrGlySerProTyrAlaAsnThrAlaGluValLeuProThrAlaAla 269
 QY 945 ATGAGCTCTCACTGATCCCTCGCCTGAACCTGGCTGTAGTCTTTTCCCA----- 998
 DB 270 AlaAlaAlaGlyLeuLeuGlyHisAlaAsnLeuAlaGlyValAlaAlaPheProAlaVal 289

QY 998 ----- 998
 DB 290 LeuSerGlyPheThrGlyAsnAspLeuValAlaIleThrSerAlaLeuAsnThrLeuAla 309
 QY 999 -----GTTTCATCCAGC 1010
 DB 310 SerTyrGlyTyrAsnLeuAsnThrLeuGlyLeuGlyLeuSerGlnAlaAlaThrGly 329
 QY 1011 GCAGTCCCGCGCTCCAGCAGCGTTACTCGGGGTGCTCCTATAGTCTCTTATGAG 1070
 DB 330 AlaLeuAlaAlaAlaAlaAlaSerAlaAsnProAlaAlaAlaAlaAsnLeuAla 349
 QY 1071 GCTCCGAGCAGGAGATGTCAGTGTATTCCCGCCAGGAGTGGCGGCATCATC 1130
 DB 350 ThrTyrAlaSerGlu-----AlaSerAlaSerGlySerThrAla 362
 QY 1131 GGCAGAGAGGGGCAGCACATCAACAGCTCTCCGGTTTGGCAGCGCTCCATCAAGATT 1190
 DB 363 GlyGlyThrAla-----GlyThrPheAlaLeuGlySerLeuAlaAla 376
 QY 1191 GCACACCCGAAACACCTGACTCCAAAGTTGATGTTATCATCATCTGGACCGCCAG 1250
 DB 377 Ala-----ThrAlaAlaThrAsn 382
 QY 1251 GCCCAATTCAAGGCTCAGGGAAGATCTATGCAAACTC-----AAGGAGGAGAAC 1301
 DB 383 GlyTyrPheGlyAlaAlaSerProLeuAlaAlaSerAlaIleLeuGlyThrGluLysSer 402
 QY 1302 TTTTGTGTCACAGGAGGAGTGAAGCTGGAGACCCACATACATCGTGTCCAGCATCAGCA 1361
 DB 403 ThrAspGlySerLysAspValValGlu-----IleAlaValProGluAsnLeu 418
 QY 1362 GCTGCCCGGTCATTGGCAAGGTGAAACGGTGAACGAGTTGCAGATTTTGACGCA 1421
 DB 419 ValGlyAlaIleLeuGlyLysGlyLysThrLeuValGluTyrGlnGluLeuThrGly 438
 QY 1422 GCTGAGGTGGTAGTACCAGA-----GACACAGCCCTGTAGTGAACACGAGCATC 1475
 DB 439 AlaArgIleGlnIleSerLysLysGlyGluPheValProGlyThrArgAsnArgLys--- 457
 QY 1476 GTGAAATATCATCGGACATTTCTATCCAGTCAGATGGCTCAA 1517
 DB 458 ValThrIleThrGlyThrProAlaAlaThrGlnAlaAlaGln 471

RESULT 11
 NOA1_HUMAN
 ID NOA1_HUMAN STANDARD; PRT; 510 AA.
 AC P51513;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE RNA-binding protein Nova-1 (Neuro-oncological ventral antigen 1)
 DE (Onconeural ventral antigen-1) (Paraneoplastic R antigen) (Ventral neuron-specific protein 1).
 GN NOVA1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
 RC TISSUE=Cerebellum, and Hippocampus;
 RX MEDLINE=94000830; PubMed=8398153;
 RA Buckanovich R.J., Posner J.B., Darnell R.B.;
 RT "Nova, the paraneoplastic R antigen, is homologous to an RNA-binding protein and is specifically expressed in the developing motor system.";
 RT system.";
 RL Neuron 11:657-672(1993).
 RP [2]
 RN SEQUENCE OF 1-34 FROM N.A.
 RC TISSUE=Fetal brain;
 RA Dmitrenko V.V., Garifulin O.M., Shostak K.A., Smikodub A.I.,


```

Db      138 LeuProAsnSerThrGluArgAlaIleThrIleAlaGlyValProGlnSerValThrGlu 157
QY      888 GCCGAGCAGGAAATA-----ATGAAGAAAGTTCCGGAGGCCCTATCAGAATGAT 935
Db      158 CysValLysGlnIleCysLeuValMetLeuGluThrLeuSerGlnSerProGlnGlyArg 177
QY      936 GTGGCTGCCATGAGCTTCACCTGATCCCTGGCTGCAACCTGGCT----- 980
Db      178 ValMetThrIleProTyrGlnProMetProAlaSerSerProValIleCysAlaGlyGly 197
QY      981 -----GCTGTAGTCTTTTCCAGCTTCATCCAGCGCATGCCCGCG 1022
Db      198 GlnAspArgCysSerAspAlaValGly---TyrProHisAlaThrHisAspLeuGluGly 216
QY      1023 CTTCCC----- 1028
Db      217 ProProLeuAspAlaTyrSerIleGlnGlyGlnHisThrIleSerProLeuAspLeuAla 236
QY      1028 ----- 1028
Db      237 LysLeuAsnGlnValAlaArgGlnGlnSerHisPheAlaMetMetHisGlyGlyThrGly 256
QY      1029 ---AGCAGCGTTACTGGCGCTGCTCCC-----TATAGCTCTTTATCAGGCT 1073
Db      257 PheAlaGlyIleAspSerSerProGluValLysGlyTyrTrpAlaSerLeuAspAla 276
QY      1074 CCGCAGCAGGAGATGGTGCAGCTTTATCCCGCCCGCAGGCGGCGCATCATCGGC 1133
Db      277 SerThrGlnThrThrHisGluLeuThrIleProAsnAsnLeuIleGlyCysIleIleGly 296
QY      1134 AAGNAGGGGAGCAGCATCAACACTCTCCGGTTTCCGCGCTCCATCAAGATTGCA 1193
Db      297 ArgGlnGlyAlaAsnIleAsnGluIleAargGlnMetSerGlyAlaGlnIleLysIleAla 316
QY      1194 CCACCCGAAACACCTGACTCCAAAGTTCGTATGTTATCATCACTGCGCCGAGGCC 1253
Db      317 AsnPro---ValGluGlySerSerGlyArgGlnValThrIleThrGlySerAlaAlaSer 335
QY      1254 CAATTCAAGGCTCAGGGAAGATATGTCGCAAACTCAAGGAGGAGCAACTCTTTGGT 1310
Db      336 IleSerLeuAlaGlnTyrLeuIleAsnAlaArgLeuSerSerGluLysGlyMetGly 354

RESULT 10
NOAL_MOUSE
ID NOAL_MOUSE STANDARD; PRT; 493 AA.
AC Q3XNG; Q8CB9;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE RNA-binding protein Nova-1 (Neuro-oncological ventral antigen 1)
DE (Ventral neuron-specific protein 1) (Fragment).
GN NOVA1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=CBA/J; TISSUE=Brain;
RX MEDLINE=20458856; PubMed=11003693;
RA Ward-Bailey P.F., Wood B., Johnson K.R., Bronson R.T., Donahue L.R.,
RA Davison M.T.;
RT "Neuromuscular ataxia: a new spontaneous mutation in the mouse.";
RL Mamm. Genome 11:820-823 (2000).
RN [2]
RN SEQUENCE OF 250-493 FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=22354683; PubMed=1246851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,

```

```

RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla B., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gasterlund S., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustinchin S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wallestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayaehizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RT Nature 420:563-573 (2002).
RN [3]
RN FUNCTION.
RP MEDLINE=20182810; PubMed=10719891;
RX Jensen K.B., Dredge B.K., Stefani G., Zhong R., Buckanovich R.J.,
RA Okano H.J., Yang Y.Y.-L., Darnell R.B.;
RT "Nova-1 regulates neuron-specific alternative splicing and is
RL essential for neuronal viability.";
RL Neuron 25:359-371 (2000).
RN [4]
RN FUNCTION.
RX MEDLINE=22692957; PubMed=12808107;
RA Dredge B.K., Darnell R.B.;
RT "Nova regulates GABA(A) receptor gamma2 alternative splicing via a
RT distal downstream UCAU-rich intronic splicing enhancer.";
RT Mol. Cell. Biol. 23:4687-4700 (2003).
CC -1- FUNCTION: Functions to regulate alternative splicing in neurons by
CC binding pre-mRNA in a sequence-specific manner to activate exon
CC inclusion. It binds specifically to the sequence UCAU. Most
CC likely acts to activate the inclusion of exon E3a in the glycine
CC receptor alpha-2 chain and of exon E9 in gamma-aminobutyric-acid
CC receptor gamma-2 subunit via a distal downstream UCAU-rich
CC intronic splicing enhancer.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DISEASE: Defects in NOVA1 leads to neuronal death in spinal and
CC brainstem neurons.
CC -1- SIMILARITY: Contains 3 KH domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF232828; AAF35907.1; -.
CC EMBL; AK047565; BAC33088.1; -.
CC MGD; MGI:104297; Novol.
CC GO; GO:0005634; C:nucleus; IC.
CC GO; GO:0003729; F:mRNA binding; IDA.
CC GO; GO:0006371; P:mRNA splicing; IDA.
CC InterPro; IPR004087; KH_dom.
CC InterPro; IPR004086; KH_type_1.
CC Pfam; PF00013; KH; 3.
CC SMART; SM00322; KH; 3.
CC PROSITE; PS00084; KH TYPE 1; 3.
KW Antigen; Nuclear protein; RNA-binding; Repeat.
PT NON_TER 1 1

```


RA Yang Y.Y., Yin G.L., Darnell R.B.;
 RT "The neuronal RNA-binding protein Nova-2 is implicated as the
 RT autoantigen targeted in POMA patients with dementia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:13254-13259(1998).
 [3]
 RP SEQUENCE OF 29-492 FROM N.A.
 RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,
 RA Burkhardt-Schultz K.J., Gordon L., Dias J., Ramirez M., Scilwagen S.,
 RA Phan H., Velasco N., Do L., Regala W., Terry A., Gaines J.,
 RA Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,
 RA Attix C., Andreise T., Trankheim M., Amico-Keller G., Coesfield J.,
 RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,
 RA Arellano A., Brower A., Sanders C., Ow D., Nolan M., Trong S.,
 RA Kobayashi A., Olsen A.S., Carrano A.V.;
 RA "Sequence analysis of a 1.9 Mb region in 19q13.2 between APOE and
 RT D19S412.";
 RT Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
 [4]
 RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 405-480.
 RX MEDLINE=99148126; PubMed=10368286;
 RA Lewis H.A., Chen H., Edo C., Buckanovich R.J., Yang Y.Y.-L.,
 RA Musunuru K., Zhong R., Darnell R.B., Burley S.K.;
 RT "Crystal structures of Nova-1 and Nova-2 K-homology RNA-binding
 RT domains.";
 RL Structure 7:191-203(1999).
 [5]
 RN X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 405-491 IN COMPLEX WITH RNA.
 RX MEDLINE=20139984; PubMed=10676814;
 RA Lewis H.A., Musunuru K., Jensen K.B., Edo C., Chen H., Darnell R.B.,
 RA Burley S.K.;
 RT "Sequence-specific RNA binding by a Nova KH domain: implications for
 RT paraneoplastic disease and the fragile X syndrome.";
 RL Cell 100:323-332(2000).
 CC -!- FUNCTION: May regulate RNA splicing or metabolism in a specific
 CC subset of developing neurons (By similarity). Binds single strand
 CC RNA.
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- TISSUE SPECIFICITY: Brain. Expression restricted to astrocytes.
 CC -!- DOMAIN: The third KH domain (KH3) recognizes specifically 5'-
 CC UCAY-3'.
 CC -!- DISEASE: Autoantigen in the paraneoplastic opsoclonus myoclonus
 CC ataxia (POMA), a paraneoplastic neurological syndrome/disorder
 CC (PNS/P) associated with breast cancer, fallopian cancer, and SCLCa
 CC and characterized primarily by loss of inhibitory control of motor
 CC neurons in the spinal cord and brainstem. Recognized by the IgG
 CC autoantibody ANNA-2 (also called anti-R1).
 CC -!- SIMILARITY: Contains 3 KH domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U70477; AAB88661.1; ALT_INIT.
 DR EMBL; AF083898; AAC72355.1; -.
 DR EMBL; AC006540; AAD13116.1; -.
 DR PDB; 1DTJ; 18-FEB-00.
 DR PDB; 1EC6; 07-APR-00.
 DR Genew; HGNC:7887; NOVA2.
 DR MIM; 601991; -.
 DR InterPro; IPR004087; KH_dom.
 DR InterPro; IPR004088; KH_type_1.
 DR Pfam; PF00013; KH; 3.
 DR SMART; SM00322; KH; 3.
 DR PROSITE; PS50084; KH_TYPE_1; 3.
 KW Antigen; Nuclear protein; RNA-binding; Repeat; 3D-structure.
 FT DOMAIN 10 26 BIPARTITE NUCLEAR LOCALIZATION SIGNAL
 FT FT
 FT DOMAIN 32 99 KH 1.
 FT FT
 FT DOMAIN 130 196 KH 2.
 FT FT

FT	DOMAIN	232	405	ALA-RICH.
FT	DOMAIN	332	393	GLY-RICH.
FT	DOMAIN	406	473	KH 3.
FT	CONFLICT	247		A -> R (IN REF. 1).
FT	CONFLICT	265	267	PAA -> TAT (IN REF. 1).
SQ	SEQUENCE	492 AA;	49008 MW;	41B63EAF6899256B CRC64;

Alignment Scores:				
Pred. No.:	2.12e-07	Length:	492	
Score:	236.50	Matches:	118	
Percent Similarity:	37.89%	Conservative:	62	
Best Local Similarity:	24.84%	Mismatches:	185	
Query Match:	6.70%	Indels:	111	
DB:	1	Gaps:	15	

US-09-270-437D-7 (1-1946) x NOA2_HUMAN (1-492)

QY	390	TCTCGGGGTGAGCCCGCCGAGGGTCCACCT	-----GTGGCAGCGGGGGCCCGCCAGC	443
DB	8	SerArgLysArgProLeuGluThrProGluValValCysThrLysArgSerAsnThr	27	
QY	444	CAGCAGCAAGTGAGACATCCCTTCGGCTCTCGGTCTCTGGTGGCCACCAGTANGTGGTGGCCATT	503	
DB	28	GlyGluGluGlyGluTyrPheLeuLysValLeuLeuProSerTyrAlaAlaGlySerIle	47	
QY	504	ATTGGCAGAGGAGGGGCCACCATCGCAACATCACAAACACAGCCACGATCCCAAGATAGAC	563	
DB	48	IleGlyLysGlyGlyGlnThrIleValGlnLeuGlnLysGlnThrGlyAlaThrIleLys	67	
QY	564	GTGCATAGCAAGAGGAAC-----GCAGGTGCAGCTGAAAGAACCATCATGTGTGCATCC	617	
DB	68	LeuSerLysSerLysAspPheTyrProGlyThrThrGluArgValCysLeuValGlnGly	87	
QY	618	ACCCCTGAGGGCTGCTCTCGCTTGTGAAGATGATCTTGGAGATATATGCATAAAGAGGCT	677	
DB	88	ThrAlaGluAlaLeuAsnAlaValHisSerPheIleAlaGluLysVal---ArgGluIle	106	
QY	678	AAGGACACCAAAACGGCTGACGAGGTT-----	704	
DB	107	ProGlnAlaMetThrLysProGluValValAsnIleLeuGlnProGlnThrThrMetAsn	126	
QY	705	-----CCCTCGAAGATCTCGGCCATAATACTTTGTAGGGCGTCTCAT	749	
DB	127	ProAspArgAlaLysGlnAlaLysLeuIleValProAsnSerThrAlaGlyLeuIleIle	146	
QY	750	GCCAGGAGGAGCGAAGCTGAGAGGTAGGAGTAGGCAAGATACGACACAAATACCATC	809	
DB	147	GlyLysGlyGlyAlaThrValLysAlaValMetGluGlnSerGlyAlaTrpValGlnLeu	166	
QY	810	TCC-----TCGTTGCAAGACCTTACCTTTTACACCTCGAGGAGCC	851	
DB	167	SerGlnLysProGluGlyIleAsnLeuGln-----GluArgVal	179	
QY	852	AFCACCTGTAAGGGGCCCATCGAGAAATTGTCAGGGCGGACGAGCAATAATAGAGAA	911	
DB	180	ValThrValSerGlyGluProGluGlnValHisLysAlaValSerAlaIleValGlnLys	199	
QY	912	GTTCCGGAG-----GCCTATGAGAT-----	932	
DB	200	ValGlnGluAspProGlnSerSerCysLeuAsnIleSerTyrAlaAsnValAlaGly	219	
QY	933	-----GATGTG-----	938	
DB	220	ProValAlaAsnSerAsnProThrGlySerProTyrAlaSerProAlaAspValLeuPro	239	
QY	939	-----GCTGCCATGAGCTCTACCTGATCCCTGGCGCTGAACCTGGCT	980	
DB	240	AlaAlaAlaAlaAlaSerAlaAlaAlaSerGlyLeuLeuGlyProAlaGlyLeuAla	259	
QY	981	GCTGTAGGTCTTTCCAGCTTC-----ATCCAGCGC	1012	
DB	260	GlyValGlyAlaPheProAlaAlaLeuProAlaPheSerGlyThrAspLeuLeuAlaIle	279	

QY 332 CATCCCGATGAGCAGATAGCACAGGACCTGAGAAATGGGGCGCGAGGGGGCTTTGGGCTC 391
Db |||||
171 ---ProArg- |||||
QY 392 TCGGGGTGAGCCCGCCAGGGCTCACCTGTGGC- - - - -AGCGGGGGGCCCGCAGC 439
Db |||||
177 ---ArgArgProProProAlaLeuThrGlyGlyAlaLeuProSerAlaLeuProPro 195
QY 440 CAAGCAGCAGCAGTGGACATCCCTCTCGG- - - - -CTCCTGGT 478
Db |||||
196 GlnLeuGlyProMetHisProProPro-ArgSerThrValThrGluGluTyrArgVa 215
QY 479 GCCCACCAGTATGTGGTGCATTTATTGGCAAGGAGGGGCCACCATCCGCAACATCAC 538
Db |||||
215 lProAspGlyMetValGlyLeuIleIleGlyArgGlyGlyGlnIleAsnIleGly 235
QY 539 AAAACACACCCAGTCCCAAGATAGACGTGCATAGGAGGAGAACCGAGTGCAGCTGA 598
Db |||||
235 nGlnAspSerGlyCysIysValGlnIle- - - - -SerProAspSerGlyGlyLeuProGluAr 254
QY 599 AGCCATCAGTGTGACCTCCACCTCGAGGGCTGCTCTCCGCTTGTAAAGATGATCTTGA 658
Db |||||
254 gSerValSerLeuThrGlySerProGluAlaValGlnIleLysAlaLysLeuMetLeuAspAs 274
QY 659 GATTATG- - - - -CATAAAGAGGCTAAG- - -GA 682
Db |||||
274 pIleValSerArgGlyArgGlyGlyProGlyGlnPheHisAspTyrAlaAsnGlyGl 294
QY 683 CACCAAAACGGTGCACGAGGTCCCTCGAAGATCTCGCCCATATTAATTTGTAGGGCG 742
Db |||||
294 nAsnGlyThrValGlnGlu- - - - -IleMetIleProAlaGlyLysAlaGlyLe 310
QY 743 TCTCATTTGCAAGGAAGGACGGAACCTGAAGAGGTAGACAGATACCCAGAGCAAAAT 802
Db |||||
310 uValIleGlyGlyGlyGlyGluThrIleLysGlnLeuGlnGluArgAlaGlyValIysMe 330
QY 803 CACATC- - - - -TCCTCGTTGCAAGACCTTACCTTTACAACCTCGAGAGGACCATCAC 856
Db |||||
330 tIlePheIleGlnAspGlySerGlnAsnThrAsnValAspIysProLeuArgIleIleGl 350
QY 857 TGTGAAGGGGCCATCCAGATTTGTCAGGGCCGACGAGCAAGAAATATGAAGATTCG 916
Db |||||
350 yAspProTyrLysValGlnGlnAlaCysGluMetValMetAspIleLeuArgGluArgAs 370
QY 917 GGAGGCTATCAGAAATGATGTGGCTGCATCAGCTCTCACCTGATCCCTGGCTGACCT 976
Db |||||
370 pGlnGlyGlyPheGlyAspArgAsnGluTyrGlySerArgIleGlyGlyIleAsp- - 389
QY 977 GGCTGCTGTAGTCTTTCCAGCTTCATCCAGGGCAGTCCCGCCGCTCCCGACGCGT 1036
Db |||||
390 - - - - -ValPro- - - - - 391
QY 1037 TACTGGGGCTGCTCCTATAGCTCTTTATGCAGGCTCCCGAGCAGAGATGGTGAGGT 1096
Db |||||
391 - - - - - 391
QY 1097 GTTTATCCCGCCAGCAGTGGCGGCATCATCGCAAGAGGGCGAGCAGCATCAACA 1156
Db |||||
392 - - - - -ValProArgHisSerValGlyValValIleGlyArgSerGlyGluMetIleLysLy 410
QY 1157 GCTCTCCGGTTTGGCAGCGCTCCATC- - - - - 1184
Db |||||
410 sIleGlnAsnAspAlaGlyValArgIleGlnPheLysGlnAspAspGlyThrGlyProGl 430
QY 1185 -AGATTGCA- - - - -CCACCCGAAACACCTGACTCCCAAAATTCGTATGGTT- - 1229
Db |||||
430 uLysIleAlaHisIleMetGlyProProGluArgCysGluHisAlaAlaArgIleIleAs 450
QY 1230 - - - - -ATCATCATGTCAGCGCCAGAGGCCCAATT- - - - - 1258
Db |||||
450 nAspLeuLeuGlnSerLeuArgSerGlyProPro- -GlyProGlyHisGlyMetPro 469

QY 1259 - - - - -CAAGGCTCAGGAAGAATCTATGGCAAACTCAAGAGGAGAACTT 1303
Db |||||
470 ProGlyGlyArgGlyArgGlyArgGlyGlnGlyIleTrpGlyProProGlyGlyGluMet 489
QY 1304 CTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCACATACGTGTGCCAGCATCAGCAGC 1363
Db |||||
490 ThrPhe- - - - -Ser-IleProThrHisLysCy 498
QY 1364 TGGCGGGTTCATTGGCAAGGTGAAAAACGGTGAACAGAGTTGCAGATTTTCAGCGCAGC 1423
Db |||||
498 sGlyLeuValIleGlyArgGlyGlyGluAsnValLysAlaIleAsnGlnArgGlyAl 518
QY 1424 TGAGGTGTAGTACCAAGAGACACCCCT- - - - -GATGAGAACACGACGATCATCGT 1477
Db |||||
518 aPheValGluIleSerArgGlnLeuProProAsnGlyAspProAsnPheLysLeuPheI 538
QY 1478 GAAATCATCGGACATTCTATGCCAGTCAGATGGCTCAACGGGAAGATCCGAGACATCCT 1537
Db |||||
538 e- - - - -IleArgLysSerPr 543
QY 1538 GGCCCGAGTTAAGCAGCAGCATCAGAGGAGGACAGAGTAACAGGCCCCAGCGCAGGAA 1597
Db |||||
543 oGlnGlnIleGluHisAlaLysGln- - - - -ProIleGluGluLy 556
QY 1598 GTGACGAGCCCTCCCTGCTCCCTTNGAGTCCAGGACAAACACGGGCGAGAAATCGAGATG 1657
Db |||||
556 sIleGluGlyProLeuCysPro- - - - - 563
QY 1658 TGCTCTCCCGCGCAGGCTGAGATGAGTGGGAATCCGGGACACNTGGCGGCTGTAG 1717
Db |||||
564 - - - - -ValGlyProGlyProGl 569
QY 1718 ATCAGGTTTCCCATTTGATTGAGAAAGATGTTCCAGTGAGGAACCTGATCTTCAGCC 1777
Db |||||
569 yProGlyGlyProPro- - - - - 574
QY 1778 CCAACACCCACCAATGGCCCAACTGNTGCCCTCGGGGTGTCAGAAATTTAGC 1837
Db |||||
575 - - - - -GlyProAlaGly- - - - -Pr 579
QY 1838 GCAAGGCATTTTAAACGTGATTTTAAAGAGCT- - - - -CT 1876
Db |||||
579 oMetGlyProPheAsnProGlyProPheAsnGlnGlyProProGlyGlyProProHi 599
QY 1877 CAGAGGCCACCAAGAGGTGG- - - - -ATCACACCTCAGTGGAA 1916
Db |||||
599 sGlnTyrProProGlnGlyTrpGlyAsnThrTyrProGlnTrpGln 614

RESULT 6
NOA2 HUMAN STANDARD; PRT; 492 AA.
ID NOA2 HUMAN Q9UNW9; O43267; Q9UEA1;
AC Q9UNW9; O43267; Q9UEA1;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE RNA-binding protein Nova-2 (Neuro-oncological ventral antigen 2)
DE (Astrocytic NOVA1-like RNA-binding protein).
GN NOVA2 OR ANOVA OR NOVA3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20137319; PubMed=10735272;
RA Ueki K., Ramasamy S., Billings S.J., Mohrenweiser H.W., Louis D.N.;
RT "ANOVA, a putative astrocytic RNA binding protein gene that maps to
RL chromosome 19q13.3.";
RL Neurogenetics 1:31-36 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99007301; PubMed=9789075;

QY 1000 CTTTCATCCAGCGAGTCCCGCGCTCCAGAGCGTTACTGGGCTGCTCCCTATAGCT 1059
 Db 320 -----SerArgValGlyGly----- 325
 QY 1060 CTTTATCAGGCTCCCGAGGAGATGGTCAGGTGTATTATCCCGCCAGGAGTGG 1119
 Db 326 -----IleAspValProValProArgHisSerValG 336
 QY 1120 GCGCATCATCGCAAGAGGGCGACACATCAACAGCTCTCCCGTTTCCGACGCGCT 1179
 Db 336 lyValValIleGlyArgSerGlyGluMetIleLysLysIleGlnAsnAspAlaGlyVala 356
 QY 1180 CCATC-----AAGATTCCA-----C 1194
 Db 356 rgIleGlnPheLysGlnAspAspGlyThrGlyProGluLysIleAlaHisIleMetGlyP 376
 QY 1195 CACCCGAAACCTCACTCCAAAGTTCGTATGGT-----ATCATCA 1236
 Db 376 roProAspArgCysGluHisAlaAlaArgIleIleAsnAspLeuGlnSerLeuArgS 396
 QY 1237 CTGGACCCCGAGGCCCAATT-----CAAGG 1263
 Db 396 erGlyProPro--GlyProGlyAlaProGlyMetProProGlyGlyArgGlyArgG 415
 QY 1264 CTCAGGGAATCATATGGCAACTCAAGGAGGAACTTTCTTGGTCCAGGAGGAG 1323
 Db 415 yArgGlyGlnGlyAsnTrpGlyProGlyGlyGluMetThrPhe----- 430
 QY 1324 TGAAGCTGAGGACACATACCTGTGCGCAGCATCAGCAGCTGGCGGGTCATTGCCAAG 1383
 Db 431 -----Ser-IleProThrHisLysCysGlyLeuValIleGlyArgG 444
 QY 1384 GTGGAATAACCGTGAACGATTCAGAAATTCACGCGAGCTAGGTGGTAGTACCAAG 1443
 Db 444 lyGlyGluAsnValLysAlaIleAsnGlnInThrGlyAlaPheValGluIleSerArgG 464
 QY 1444 ACCAGACCCCT-----GATGAGACGACCATCGTTCATCGT-----AAAA 1482
 Db 464 lnLeuProProAsnGlyAspProAsnPheLysLeuPheValIleArgGlySerProGlnG 484
 QY 1483 TCATCGGACATTTCTATGCGCAGTCAGATGGCTCAACGGAAGATCCGACAGATCCTGGCC 1542
 Db 484 lnIleAspHis-----AlaLysGlnLeuIleGluLysIleGlu----- 497
 QY 1543 AGGTAAAGCAGCATCATCAAGAGGACAGAGTAACGAGGCCCGCCAGGACGAGGAGTAGC 1602
 Db 497 ----- 497
 QY 1603 CAGCCCTCCCTGTCCTTNGAGTCCAGGACACACGCGGACAAATCAGAGTGTGCTC 1662
 Db 498 --GlyProLeuCysPro----- 502
 QY 1663 TCCCCGCGAGCGCTGAGATGAGTGGGAATCCGGGACACNTGGCGGGCTGTAGATCAG 1722
 Db 503 -----ValGlyProGly----- 506
 QY 1723 GTTGTGCCACTTGATTGAGAAAGATGTTCCAGTGAGGAACCTGTATCTNTACGCCCAAA 1782
 Db 506 ----- 506
 QY 1783 CACCCACCAATTTGGCCCAACACTGNTGCCCTCGGGGTGTGAGAAATTTAGGCAAG 1842
 Db 507 -----ProGlyGlyProGlyProAlaGlyProMetG 517
 QY 1843 GCACCTTTTAAACGTGATTGTTTAAAGAAGCT----- 1874
 Db 517 lyProPheHisProGlyProPheAsnGlnGlyProGlyAlaProProHisAlaGlyG 537
 QY 1875 -----CTCCAGGCCCAACCAAGAGGTGG-----ATCACCTCACTGAGTGGAA 1916
 Db 537 lyProProHisGlnTyrProProGlnGlyTrpGlyAsnThrTyrProGlnTrpGln 556

RESULT 5

ZBP2_CHICK
 ID ZBP2_CHICK STANDARD; PRT; 769 AA.
 AC Q8UVD9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Zipcode-binding protein 2.
 GN ZPB2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A., SEQUENCE OF 104-119; 242-267; 679-685; 705-711 AND
 RP 743-766, FUNCTION, DEVELOPMENTAL STAGE, AND SUBCELLULAR LOCATION.
 RX MEDLINE=21640447; PubMed=11781334;
 RA Gu W., Pan F., Zhang H., Bassell G.J., Singer R.H.;
 RT "A predominantly nuclear protein affecting cytoplasmic localization of
 RT beta-actin mRNA in fibroblasts and neurons.";
 RL J. Cell Biol. 156:41-52(2002).
 CC -!- FUNCTION: Binds to a 54-nucleotide localization signal (the
 CC zipcode) found in the 3' untranslated region of beta-actin mRNA
 CC and may play a role in its localization.
 CC -!- SUBCELLULAR LOCATION: Predominantly nuclear. Also detected in the
 CC cytoplasm.
 CC -!- DEVELOPMENTAL STAGE: The highest expression is found in 6-d
 CC embryos, is reduced to 30% before hatching and remains stable
 CC thereafter.
 CC -!- SIMILARITY: Contains 4 KH domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF461020; AAL66365.1; -;
 DR GO; GO:0005737; C:cytoplasm; IDA.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR GO; GO:0003730; F:mRNA 3' UTR binding; IDA.
 DR GO; GO:0008298; P:mRNA localization, intracellular; NAS.
 DR InterPro; IPR004087; KH_dom.
 DR InterPro; IPR004088; KH_type_1.
 DR Pfam; PF00013; KH; 4.
 DR SMART; SM00322; KH; 4.
 DR PROSITE; PS50084; KH TYPE 1; 4.
 DR Transport; mRNA transport; Nuclear protein; RNA-binding; Repeat.
 KW DOMAIN 208 272 KH 1.
 FT DOMAIN 296 362 KH 2.
 FT DOMAIN 385 449 KH 3.
 FT DOMAIN 486 553 KH 4.
 SQ SEQUENCE 769 AA; 80644 MW; 875A6C83529969EB CRC64;
 Alignment Scores:
 Pred. No.: 2,05e-07 Length: 769
 Score: 237.50 Matches: 135
 Percent Similarity: 32.29% Conservative: 71
 Best Local Similarity: 21.16% Mismatches: 207
 Query Match: 6.73% Indels: 225
 DB: 1 Gaps: 25
 US-09-270-437D-7 (1-1946) x ZBP2_CHICK (1-769)
 QY 212 GAACACCGAGGTGACGACGCGAGTGGTGAATGTACCTATTCCAAACCGGAGCAGACAG 271
 Db 135 GluSerLysLysLeuAlaAlaGlnGlyAspCysGluTyrGlyProArgGlyGlyAlaPro 154
 QY 272 GCAAGCCATCATGAAGCTGAATGGCCACCGAGTTGGAGAACCATGCTCCCTGAAGTCTCTTA 331
 Db 155 AlaAlaProGluArgSerGlyProValGlyAspProGlyProGlyProGlyProGlyPro 170

QY 1813 CCCTCGGGTGTGAGAAATTTAGCGGACGACCTTTAAACGTGGATTGTTTAAAGAG 1872
 Db 560 nThAsnGlyGlnGlyAspGlnGlnAla----- 569
 QY 1873 CTCTCCAGGCCCCACCAAGAGGGTGGATCACACC-----TC 1908
 Db 570 ---ProAlaProAlaGlyGlnValAspTyrThrLysAlaTrpGluGluTyrTyrLysLys 588
 QY 1909 AGTCGGAGAAATAAATAATTCCTTCAGGT 1939
 Db 588 sMetGlyGlnAlaValProAlaProAlaGly 598

RESULT 3
 FUBI_HUMAN
 ID FUBI_HUMAN STANDARD; PRT; 643 AA.
 AC Q96AE4; Q12828;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Far upstream element binding protein 1 (FUSE binding protein 1) (FBP)
 DE (DNA helicase V) (HDH V).
 GN FUBP1.
 OS Homo sapiens (Human).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1), SEQUENCE OF 191-193; 203-206;
 RP 272-279; 284-290; 300-314; 321-328; 394-397; 409-411; 430-438 AND
 RP 440-443, AND FUNCTION.
 RC TISSUE=Leukemia;
 RX MEDLINE=94170991; PubMed=9125259;
 RA Duncan R., Bazar L., Michelotti G., Tomonaga T., Krutzsch H.,
 RA Avigan M., Levens D.;
 RT "A sequence-specific, single-strand binding protein activates the far
 RT upstream element of c-myc and defines a new DNA-binding motif";
 RL Genes Dev. 8:465-480(1994).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Reingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Diatchenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.H., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE OF 45-64; 133-145; 271-283; 308-321; 379-386; 414-424 AND
 RP 430-439.
 RX MEDLINE=21127960; PubMed=11222755;
 RA Vindigni A., Ocham A., Triolo G., Falaschi A.;
 RT "Identification of human DNA helicase V with the far upstream
 RT element-binding protein";
 RL Nucleic Acids Res. 29:1061-1067(2001).
 RN [4]
 RP PARTIAL SEQUENCE, AND MASS SPECTROSCOPY.
 RX MEDLINE=22166132; PubMed=12176931;

RA Rappsilber J., Ryder U., Lamond A.I., Mann M.;
 RT "Large-scale proteomic analysis of the human spliceosome.";
 RL Genome Res. 12:1231-1245(2002).
 RN [5]
 RP INTERACTION WITH SIAHBPI.
 RX MEDLINE=20337922; PubMed=10882074;
 RA Liu J., He L., Collins I., Ge H., Libutti D., Li J., Egly J.-M.,
 RA Levens D.;
 RT "The FBP interacting repressor targets TFIID to inhibit activated
 RT transcription";
 RL Mol. Cell 5:331-341(2000).
 RN [6]
 RP INTERACTION WITH JTV1, UBIQUITINATION, AND PROTEASOME-MEDIATED
 RP DEGRADATION.
 RX MEDLINE=22716800; PubMed=12819782;
 RA Kim M.-J., Park B.-J., Kang Y.-S., Kim H.J., Park J.-H., Kang J.W.,
 RA Lee S.W., Han J.M., Lee H.-W., Kim S.;
 RT "Downregulation of FUSE-binding protein and c-myc by TRNA synthetase
 RT cofactor p38 is required for lung cell differentiation";
 RL Nat. Genet. 34:330-336(2003).
 RN [7]
 RP STRUCTURE BY NMR OF 278-447 IN COMPLEX WITH SINGLE STRANDED FUSE DNA.
 RX MEDLINE=21864720; PubMed=11875576;
 RA Braddock D.T., Louis J.M., Baber J.L., Levens D., Clore G.M.;
 RT "Structure and dynamics of KH domains from FBP bound to
 RT single-stranded DNA";
 RL Nature 415:1051-1056(2002).
 CC -!- FUNCTION: Regulates MYC expression by binding to a single-stranded
 CC far-upstream element (FUSE) upstream of the MYC promoter. May act
 CC both as activator and repressor of transcription.
 CC -!- SUBUNIT: Interacts with SIAHBPI/PIR and JTV1.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC Name=2;
 CC IsoId=Q96AE4-1; Sequence=Displayed;
 CC IsoId=Q96AE4-2; Sequence=VSP_008321;
 CC Note=No experimental confirmation available;
 CC -!- PTM: Ubiquitinated. This targets the protein for proteasome-
 CC mediated degradation.
 CC -!- SIMILARITY: Contains 4 KH domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U05040; AAA17976.2; -;
 CC EMBL; BC017247; AAH17247.1; -;
 CC PDB; 1J4W; 06-MAR-02
 CC Genew; HGNC:4004; FUBP1.
 CC MIM; 603444; -;
 CC InterPro; IPR008160; Collagen.
 CC InterPro; IPR004087; KH dom.
 CC InterPro; IPR004088; KH_type_1.
 CC Pfam; PF00013; KH; 4.
 CC SMART; SM00322; KH; 4.
 CC PROSITE; PS50084; KH TYPE 1; 4.
 CC Transcription regulation; Trans-acting factor; Nuclear protein;
 CC DNA-binding; Repeat; Ubl conjugation; Alternative splicing;
 CC 3D-structure.
 CC 3D-structure. 99 163 KH 1.
 CC FT DOMAIN 184 250 KH 2.
 CC FT DOMAIN 274 338 KH 3.
 CC FT DOMAIN 375 442 KH 4.
 CC FT DOMAIN 13 26 GLY-RICH.
 CC FT DOMAIN 348 395 GLY-RICH.
 CC FT DOMAIN 449 559 PRO-RICH.
 CC FT VARSPIC 642 643 GQ -> CRDPASIELAL (in isoform 2).
 CC FT

[illegible]

```

QY 1419 GCAGCTGAGTGGTAGTACCAAGAGACGAGCCCGCTGATGAGACGACCAGGTTCATCGT 1478
Db 385 GlyAlaHisValGluLeuGlnArgAsnProProAsnSerAsp----- 399
QY 1479 AAAATCATCGACATTTCTATGCGCATGATGCTCAACGGAAGATCCGAGACATCGT 1538
Db 400 -----ProAsnLeuArgArgPheThrLeuArgGlyValPro 411
QY 1539 GCCCAGGTTAAGCAGCAGCATCAGAAG----- 1565
Db 412 GlnGlnIleGluValAlaArgGlnLeuLeuAspGluLysValGlyGlyThrAsnLeuGly 431
QY 1566 -----GGACAGAGTAACCGCCCGCAGGACGAGGAGTGCACCGCCCT 1610
Db 432 AlaProGlyAlaPheGlyGlnSerProPheSerGln-----Pro 444
QY 1611 CCCTGTCCCTTNGAGTCAGACAAACGCGGCAAGAAATGAGAGTGTGCTCTCCCGGC 1670
Db 445 ProAlaProProHisGlnAsnThrPheProProArgSerSerGlyCys-PheProAsnMe 464
QY 1671 AGGCCTGAGAAATGAGTGGGAATCCG----- 1695
Db 464 talAlaLysValAsnGlyAsnProHisSerThrProValSerGlyProProAlaPheLe 484
QY 1696 -----GGACACATGGCGCGGTGTAGATCAGGTTGCCACCTTGATTGAGA 1742
Db 484 uThrGlnGlyTrpGlySerThrTyrglnAlaTrpGlnGln-----ProThrGlnGlnVa 502
QY 1743 AAGATGTTCCAGTACGAGAACCTGATCTNTCAGCCCAACACCCACCAATGG 1797
Db 502 lProSerGlnGlnSerGlnProGlnSerSerGlnProAsnTyrSerLysAlaTrp 520

RESULT 2
FUB1_MOUSE
ID FUB1_MOUSE STANDARD; PRT; 651 AA.
AC Q91WJ8; O8COV8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Far upstream element binding protein 1 (FUSE binding protein 1) (FBP).
GN FUBP1 OR D3BRTD30B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Retina;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Hahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RN SEQUENCE OF 1-286 FROM N.A. (ISOFORM 2).
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;

```

```

RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Mateuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frough J.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustinch S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltacea K., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Perte G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.D., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yangisawa M., Yang I., Yang L.,
RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
CC -1- FUNCTION: Regulates MYC expression by binding to a single-stranded
CC far-upstream element (FUSE) upstream of the MYC promoter. May act
CC both as activator and repressor of transcription (By similarity).
CC -1- SUBUNIT: Interacts with SIN3BP1/FIR and JTV1 (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC Name=2;
CC IsoId=Q91WJ8-1; Sequence=Displayed;
CC IsoId=Q91WJ8-2; Sequence=VSP_008322;
CC Note=No experimental confirmation available;
CC -1- PTM: Ubiquitinated. This targets the protein for proteasome-
CC mediated degradation (By similarity).
CC -1- SIMILARITY: Contains 4 KH domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; BC014763; AAH14763.1; -
DR EMBL; AK029458; BAC26457.1; -
DR MGD; MGI:1196294; D3Ertcd330e.
DR MGD; MGI:1924642; 9530027K12Rik.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH_type_1.
DR Pfam; PF00013; KH; 4.
DR SMART; SM00322; KH; 4.
DR PROSITE; PS00084; KH_TYPE 1; 4.
DR Transcription regulation; Trans-acting factor; Nuclear protein;
DR DNA-binding; Repeat; Ubl conjugation; Alternative splicing.
FT DOMAIN 96 160 KH 1.
FT DOMAIN 181 247 KH 2.
FT DOMAIN 271 335 KH 3.
FT DOMAIN 372 439 KH 4.
FT DOMAIN 16 21 POLY-GLY.
FT DOMAIN 345 392 GLY-RICH.
FT DOMAIN 446 556 PRO-RICH.
FT VARSPLOC 67 67 D -> DGSWNPSTTHWEGMPSPFKD (in isoform

```

```

RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: May interact with single-stranded DNA from the far-
CC upstream element (FUSE). May activate gene expression.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q96I24-1; Sequences=Displayed;
CC Name=2;
CC IsoId=Q96I24-2; Sequence=VSP_008323, VSP_008324;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Detected in a number of cell lines.
CC -!- SIMILARITY: Contains 4 KH domains.
CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 18.
CC -!- CAUTION: Ref.2 (AAH01325) sequence differs from that shown due to
CC a frameshift in position 527.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U69127; AAC50893.1; ALT_FRAME.
DR EMBL; BC001325; AAH01325.1; ALT_FRAME.
DR EMBL; BC007874; AAH07874.1; -.
DR Genew; HGNC:4005; FUBP3.
DR MIM; 603536; -.
DR InterPro; IPR004087; KH_dom.
DR InterPro; IPR004088; KH_type_1.
DR Pfam; PF00013; KH; 3.
DR SMART; SM00322; KH; 3.
DR PROSITE; PS50084; KH_TYPE_1; 4.
KW Transcription regulation; Trans-acting factor; Nuclear protein;
KW DNA-binding; Repeat; Alternative splicing.
FT DOMAIN 77 141 KH 1.
FT DOMAIN 162 228 KH 2.
FT DOMAIN 253 317 KH 3.
FT DOMAIN 354 421 KH 4.
FT VARSPPLIC 1 64
FT STPLVDSVYGVQKRPDDGV -> MPPI (in isoform 2).
FT /FTId=VSP_008323.
FT Missing (in isoform 2).
FT /FTId=VSP_008324.
FT V > D (IN REF. 1).
FT CONFLICT 72 72 V > D (IN REF. 1).
FT SEQUENCE 572 AA; 61640 MW; F1BE22J542BC197D CRC64;
Alignment Scores:
Pred. No.: 5-59e-09 Length: 572
Score: 262.00 Matches: 126
Percent Similarity: 37.31% Conservative: 90
Best Local Similarity: 21.76% Mismatches: 218
Query Match: 7.42% Indels: 145
DB: 1 Gaps: 21
US-09-270-437D-7 (1-1946) x FUB3_HUMAN (1-572)
QY 282 ATGAGCTGAATGGC-----CACCAGTTGGAGAACCATCCCTCGAGAGTC 326
Db MetLysAlaGluGlyPheValAspAlaLeuHisArgValArgGlnIleAlaAlaIle 33
QY 327 TCCTACATCCCGATGACGAGATAGACAGGACCTCAGAAATGGCGCCGAGGGGCTTT 386
Db AspSerIleProHisLeuAsnAsnSerThrProLeuValAspSerValTyrGlyTyr 53
QY 387 GGCTCTCGGGGTGAGCC---CGCCAGGGGTACCTGTGGCAGCGGGGGCCCGAGCAAG 443
Db GlyValGlnLysArgProLeuAspAspGlyValGlyAsnGlnLeuGlyAlaLeuValHis 73

```

```

444 CAGCAGCAAGTGGACATCCCTTCGGCTCCTGGTGCCACCCAGATATGTGGTCCCAT 503
Db GlnArgThrVal---IleThrGluGluPheLysValProAspLysMetValGlyPheIle 92
504 ATTGCGAGGAGGGGCCACCATCCGCAACATCAACAAACAGACCCAGTCCCAAGATAG 563
Db IleGlyArgGlyGluGlnIleSerArgIleGlnAlaGluSerGlyCysIleGln 112
564 GTGCATAGGAAGAGAACGCGAGGTGACGTGCAAAAAGCCATCATGTGTGCATCCAC 623
Db ---AlaSerGluSerSerGlyIleProGluArgProCysValLeuThrGlyThrPro 131
624 GAGGCTCTCTCTCGCTTGAAGATGATCTTGGAGATTATG----- 665
Db GluSerIleGluGlnAlaLysArgLeuGlyGlnIleValAspArgCysArgAsnGly 151
666 -----CATAAAGAGGCTAAGGACACCAACCAACGGCTGACGAGTTCCCTGGAATC 716
Db ---ProGlyPheHisAsnAspIleAspSerAsnSerThrIleGlnGlu-----Ile 167
717 CTGGCCCAATAAATCTTTTAGGGCGCTCATTTGGCAAGAGGACGGAACCTGAAGAG 776
Db ---LeuIleProAlaSerLysValGlyLeuValIleGlyArgGlyGlyGluThrIleLysGln 187
777 GTAGAGCAAGATACCGACACAAATACCATCTCTCTGTTGCGAGACCTTACCCTT--- 833
Db LeuGlnGluArgThrGlyValLysMetValMet-----IleGlnAspGlyProLeuPro 205
834 TACAACCTGTGAGGACCATCATCTGTGAAGGGGGCCATCGAGAATTGTTGAGGCGCCAG 893
Db ThrGlyAlaAspLysProLeuArgIleThrGlyAspAlaPheLysValGlnGlnAlaArg 225
894 CAGGAATAATGAAGAAAGTTCGGAGCGCTATGAGAATGATGTGGTCCCATAGCTCT 953
Db GluMetValLeuGluIleArgLysValGlyAspGlnAlaAspPheArgGlyValArgGly 245
954 CACCTGATCCCTGGCTGAACCTGGCTGCTAGTCTTTTCCAGCTCTTCCAGCTTCATCCAG 1013
Db AspPhe----- 247
1014 GTCCCGCGCCTCCACGACGCGTTACTGGGGCTGCTCCCTATAGCTCTTTATGCGAGCT 1073
Db ---AsnSerArgMetGlyGlyGlySer----- 255
1074 CCGGACGAGAGATGTCAGGTGTTATCCCGCCAGGACGAGTGGGGCCCATATCGGC 1133
Db ---IleGluValSerValProArgPheAlaValGlyIleValIleGly 270
1134 AAGAAGGGCGACACATCAACAGCTCTCCCGTTTGGCAGCGCTCCATCAAGATTGCA 1193
Db ---ArgAsnGlyGluMetIleLysIleGlnAsnAspAlaGlyValArgIleGlnPheLys 290
1194 CCACCCGAA-----ACACCTGACTCCAAAGTTCGTATGTTATCATCATCGACCGCA 1247
Db ProAspAspGlyIleSerProGlu-----ArgAlaAlaGlnValMetGlyProPro 307
1248 GAG-----GCCCAATTCAAG 1262
Db AspArgCysGlnHisAlaAlaHisIleIleSerGluLeuIleLeuThrAlaGlnGluArg 327
1263 -----CCTCAGGGAAGCAATCATATGCAAACTCAAGGAGGAG 1298
Db AspGlyPheGlyGlyLeuAlaAlaAlaArgGlyArg-----GlyArgGlyArgGlyAsp 345
1299 AACTTCTTTGGTCCCAAGGAGGAGTGAAGTGGAGACCCACATACGTGTGCCATCATCA 1358
Db TrpSerValGlyAlaProGlyGlyValGln---GluIleThrTyrThrValProAlaAsp 364
1359 GCAGCTGCGCGGCTCATTTGGCAAAAGTGGAAAAACGGTGAACAGTTCGAGAAATTTGAC 1418
Db LysCysGlyLeuValIleGlyLysGlyGlyGluAsnIleLysSerIleAsnGlnGlnSer 384

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 23, 2004, 10:51:55 ; Search time 17.1191 Seconds
(without alignments)
11838.037 Million cell updates/sec

Title: US-09-270-437D-7
Perfect score: 3531
Sequence: 1 gctgtacggagggtggtgg.....atttccttcagggttttaaaa 1946

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 283362

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09270437/runat 23072004 095739 27413/app query.fasta 1.5582
-DB=SwissProt 42 -QFMT=fastan -SUFFIX=rsp -MINMATCH=0.1 -IOOPCL=0 -IOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09270437 @CGN 1.1 57 @runat 23072004 095739 27413 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCKS=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	262	7.4	572	1 FUB3_HUMAN	Q96124 homo sapien
2	251	7.1	651	1 FUB1_MOUSE	Q91w78 mus musculu
3	247.5	7.0	643	1 FUB1_HUMAN	Q96ae4 homo sapien
4	244.5	6.9	721	1 FUB2_RAT	Q99pf5 rattus norv
5	237.5	6.7	769	1 ZBP2_CHICK	Q8uvd9 gallus gall
6	236.5	6.7	492	1 NOA2_HUMAN	Q9unw9 homo sapien
7	234.5	6.6	707	1 FUB2_HUMAN	Q92945 homo sapien
8	231	6.5	474	1 NOA1_RAT	Q80wa4 rattus norv
9	221.5	6.3	356	1 PCB1_HUMAN	Q153k6 homo sapien
10	221.5	6.3	493	1 NOA1_MOUSE	Q9jkn6 mus musculu
11	221.5	6.3	510	1 NOA1_HUMAN	P51513 homo sapien
12	218.5	6.2	356	1 PCB1_RABIT	Q19048 oryctolagus
13	218	6.2	1268	1 VGLN_HUMAN	Q00341 homo sapien
14	217.5	6.2	403	1 PCB4_HUMAN	P57723 homo sapien
15	216.5	6.1	606	1 Y475_ARATH	P58223 arabidopsis
16	214.5	6.1	403	1 PCB4_MOUSE	P57724 mus musculu
17	208.5	5.9	339	1 PCB3_HUMAN	P57721 homo sapien
18	207.5	5.9	339	1 PCB3_MOUSE	P57722 mus musculu

19	202.5	5.7	1270	1 VGLN_CHICK	P81021 gallus gall
20	199	5.6	365	1 PCB2_HUMAN	Q15366 homo sapien
21	197.5	5.6	463	1 ROK_HUMAN	Q07244 homo sapien
22	197.5	5.6	463	1 ROK_RABIT	O19049 oryctolagus
23	194	5.5	362	1 PCB2_MOUSE	Q61990 mus musculu
24	191	5.4	413	1 PBP2_YEAST	P38151 saccharomyc
25	182.5	5.2	1222	1 S160_YEAST	P06105 saccharomyc
26	178	5.0	381	1 YBD2_YEAST	P38199 saccharomyc
27	170.5	4.8	606	1 TDRH_HUMAN	Q9y2w6 homo sapien
28	164	4.6	784	1 YAV2_XANCV	P14728 xanthomonas
29	156.5	4.4	648	1 GLT0_WHEAT	P10387 triticum ae
30	155.5	4.4	1003	1 MBD6_HUMAN	Q96dne homo sapien
31	154	4.4	839	1 GLT5_WHEAT	P10388 triticum ae
32	153.5	4.3	1820	1 STR0_HUMAN	Q9p2p6 homo sapien
33	153.5	4.3	2716	1 OSA_DROME	Q9in94 drosophila
34	153	4.3	2805	1 MAPA_HUMAN	P78559 homo sapien
35	151	4.3	838	1 GLT4_WHEAT	P08489 triticum ae
36	150	4.2	2142	1 BAT2_HUMAN	P48634 homo sapien
37	147.5	4.2	1461	1 IE18_PRIVIF	P11675 pseudorabie
38	147	4.2	1181	1 HAIR_RAT	P37609 rattus norv
39	145	4.1	660	1 YHL1_EBV	P03181 epstein-bar
40	144	4.1	5147	1 PCLO_HUMAN	Q9y6v0 homo sapien
41	143	4.0	1336	1 W146_HUMAN	Q9c0j8 homo sapien
42	141	4.0	1729	1 TABP_HUMAN	Q9c0c2 homo sapien
43	140.5	4.0	1446	1 IE18_PRIVKA	P33479 pseudorabie
44	140	4.0	1152	1 MAP4_HUMAN	P27816 homo sapien
45	139.5	4.0	931	1 MPTA_HUMAN	Q969v6 homo sapien

ALIGNMENTS

RESULT 1
FUB3_HUMAN STANDARD; PRT; 572 AA.
ID FUB3_HUMAN STANDARD; PRT; 572 AA.
AC Q96124; Q92946; Q95VB6;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Far upstream element binding protein 3 (FUSE binding protein 3).
GN FUBP3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=97094955; PubMed=8940189;
RA Davis-Smyth T., Duncan R.C., Zheng T., Michelotti G., Levens D.;
RT "The far upstream element-binding proteins comprise an ancient family
of single-strand DNA-binding transcription factors.";
RL J. Biol. Chem. 271:31679-31687(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Cervix, and Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickens M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length

This Page Blank (uspto)

DB: 2 Gaps: 13

US-09-270-437D-7 (1-1946) x S41224 (1-396)

QY 459 ATCCCCTTCGGCTCCTCGTCCGCCACCAGTAGTGCGGTGCCATTATGGAAGGAGGG 518
::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
Db 40 ValGluleuargIleuLeuGlnSerLysAsnAlaGlyAlaValIleGlyLysGlyGly 59
::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
QY 519 GCCACCATCCGCACATCACAAAACAGACCACCGACTCCAAGATAGACGTGCATAGAAGGAG 578
||||::: ::: ::::: ::::: ::::: ||||
Db 60 LysAsnIleLysAlaLeuArgThrAspTyrAsnAlaSerValSerVal-----Pro 76
::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
QY 579 AAGCGAGTGCAGCTGAAAGCAATCAGTGTGCATCTCCACCCCTGAGGCTGCTCTCTCC 638
::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
Db 77 AspSerGlyProGluArgIleLeuSerIleSerAlaAspIleGluThrIleGlyGlu 96
::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
QY 639 GCTGTGAAGATGATCTTTGGAGATTATGCATAAAGAGGCTAAGACACCAAAACGGCTGAC 698
||| |||||::: ::::: ::::: ||||
Db 97 IleLeuLysIleIleProThrLeuGluGluHisPheLysGlyAsnAspPhe----- 114
::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
QY 699 GAGGTCTCCCTGAAGATCCTGGGCCATAAACAATTTGTAGGGCTCTCATGTGCAAGGA 758
::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
Db 115 AspCysGluLeuargIleuLeuIleHisGlnSerLeuAlaGlyGlyIleIleGlyValLys 134
::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
QY 759 GGACGGAACCTGAAGAGGTAGACAGATACCCGACGACAAATAACCACTCTCTCGTTG 818
::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
Db 135 GlyAlaLysIleLysGluLeuArgGluLysThrGlnThr-----ThrIleLysLeuPhe 152
::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
QY 819 CAAGACCTTACCTTTACACCTCGAGACGACCATCACTGTGAAGGGGCCATCGAGAA 878
::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
Db 153 GluGluCysCysProHisSerThrAspArgValValLeuIleGlyGlyArgProAspArg 172
::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
QY 879 TGTTCGAGCGCGAGCAGGAGTAATAACAGAAAGTTCCGGAGGCGCTATGAGATGATGTG 938
::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
Db 173 ValValGluCysIleLysValIleLeuAspLeulleSer-----GluSerProVal 189
::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
QY 939 GCTGCCATGAGCTCTCACCTGATCCCT----- 965
||| ||||
Db 190 LysGlyArgSerGlnProTyrAspProAsnPhetyrGluThrTyraaspTyrGlyGlyPhe 209
::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
QY 966 -----GGCTGAACTGGCTGCTGTAGGT 989
||| |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
Db 210 ThrMetMetPheAspAspArgGlyArgProHISglyPheSerMetHisAlaArgGly 229
||| |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
QY 990 CTTTTCACAGCTTCATCCAGCGAGTCCCGCGCTCCAGCAGCGTTACTGGGGTCT 1049
||| |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
Db 230 GlyPhe-----AspArgMetProGlyPro-----GlyGlyArg 241
||| |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
QY 1050 CCTTATAGCTCTTTATGAGGCTCCCGAGCAGAGATGGTGCAGGTGTTTATCCCGCC 1109
|||::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
Db 242 -ProMetPro-----GlnSerArgArgAspTyrAspAspMetSerProAr 256
|||::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
QY 1110 CAGCAGTGGGGCATCATCGGCAAGAGGGGAGCAGCATCAAACAGC---TCTCCCG 1166
||| |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
Db 256 gaRGlyProLeuProProProGlyArg-ggIyGlyArgGlyGlySerArgAlaArgas 276
||| |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
QY 1167 TTTCGACGCGCTCCATCAGATTGCAACACCGAARACCTGACTCCAAGTTCGTATG 1226
||| |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
Db 276 nLeuProLeuProPro--Pro--ProProProArgGlyGlyAspArgArg----- 291
||| |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
QY 1227 GTTATCATCTGACCGCCAGAGGCCCAATTAAGGCTCAGGGAAGAATCTATGCCAAA 1286
|||::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
Db 292 -----GlyArgProAspHisTyrAspGlyMetGlyGlyArgGlyTyrGlyArg 307
|||::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
QY 1287 CTCAGGAGAGAACTCTTTTGTTCACAGGAGGAAGTGAAGCTGGAGACCCATACGT 1346
|||::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
Db 308 GlySerPheGlyAspIleGlyGlyPro-----ValIleThrThrGlnValThr 323
|||::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
QY 1347 GTCCGACATCAGCAGCTGCGCGGTCTATTGGCAAGGTGCAAAAAACGGTGAACAGCTTG 1406
|||::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
Db 324 IleProLysAspLeuAlaGlySerIleIleGlyLysGlyGlnArgIleLysGlnIle 343
|||::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||
QY 1407 CAGAAATTTGACCGCAGCTGAGGTGGTAGTACCAAGAGACACGACCCCTGATGAGAACGAC 1466
|||::: |||||:::||||:::||||:::||||:::||||:::||||:::||||:::||||


```
Db 430 ---SerAlaProAlaGlnAsnPheGlyGlyGlnGlnGlnTrpAsnProValAlaGln 448
QY 1662 CTCCTCCGGCAGCCTGAGAAATGAGTGGGAATCCGGG----- 1697
Db 449 IleProAlaAlaAlaGlnAsnProTyrGlnValGlyGlyTyrGlnGlnAsnSerValTyr 468
QY 1698 -----ACANWGGGCGGGCTGTAGATCAGGTTTG----- 1727
Db 469 AlaGlnGlnThrAlaAlaProAlaAlaProTyrAlaAlaAlaGlyIleValGlnPro 488
QY 1728 -----CCCACTTGATTGAGAAAGATGTCAGTGAGGAACC 1763
Db 489 GlnGlnGlnValAlaTyrGlnGlnProGlnValValGlnGlnVal---GlnProAlaThr 507
QY 1764 CTGATCTNTCAGCCCAACACCCACC----- 1790
Db 508 ThrValAlaAlaThrThrThrThrProThrValAspProValThrGlyGlnAspTyrSer 527
QY 1791 ---CAATTGGCCCACTGTGTGCCCCCTCGGGGTG----- 1823
Db 528 AlaGlnTrpMetGluTyrTyrLysSerIleGlyAlaHisAspLysAlaGluAlaValGlu 547
QY 1824 -----TCAGAAATNTTAGCGCAAGGCACCTTTTAAACGTGGATTG 1862
Db 548 AlaGlnMetLysLysLysAlaGluAlaAlaAlaArgAlaValProGlyGlyLeu 567
QY 1863 TTTAAAGAGCTCTCCAGGCCCCACCAAGAGGGTGGATCACCTCAG 1910
Db 568 -----IleGlnGlnMetProMetGlyMetAlaMetProGln 579

RESULT 12
I38489
onconeural ventral antigen-1 - human
C:Species: Homo sapiens (man)
C>Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 05-Nov-1999
C:Accession: I38489
R:Buckanovich, R.J.; Posner, J.B.; Darnell, R.B.
Neuron 11, 657-672, 1993
A:Title: Nova, the Paraneoplastic Ri Antigen, is Homologous to an RNA-Binding Protein an
A:Reference number: I38489; MUID:94000830; PMID:8398153
A:Accession: I38489
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-510 <RES>
A:Cross-references: EMBL:U04840; NID:G440877; PIDN:AAA16022.1; PID:G440878
C:Genetics:
A:Gene: Nova-1

Alignment Scores:
Pred. No.: 1.75e-07 Length: 510
Score: 221.50 Matches: 111
Percent Similarity: 34.89% Conservative: 68
Best Local Similarity: 21.64% Mismatches: 173
Query Match: 6.27% Indels: 161
DB: 2 Gaps: 17

US-09-270-437D-7 (1-1946) x I38489 (1-510)
QY 363 GAGATGGGGCGCGAGGGGGCTTT-----GGCTCTCGGGGTCTAG 401
Db 9 GlnAsnGlyThrHisThrGlyValProIleAspLeuAspProAspSerArgLysArg 28
QY 402 CCCCAGGAGGCTCAGCTGTGGAGGGGGGGCCCGCCAGCAGCAAGTGGAC--- 458
Db 29 ProLeuGluAlaProProGluAlaGlySerThrLysArgThrAsnThrGlyGluAspGly 48
QY 459 ----ATCCCCCTTGGCTCTCGTGGTCCCGCCAGCATGTGTGGTCCATTATTGGCAGGAG 515
Db 49 GlnTyrPheLeuLysValLeuIleProSerTyrAlaAlaGlySerIleGlyLysGly 68
QY 516 GGGGGCCACCATCCGCACATCACAACACAGACCAGTCCCAAGATAGACGTGATAGAG 575
```

```
Db 69 GlyGlnThrIleValGlnLeuGlnLysGlnThrGlyAlaThrIleLysLeuSerLysLeu 88
QY 576 GAGAACGCA-----GGTCAGGTGAAAGGCCATCAGTGTGACCTCCACC 620
Db 89 SerLysSerLysAspPheTyrProGlyThrThrGluArgValCysLeuIleGlnGlyThr 108
QY 621 CTGAGGGCTGCTCTCCGCT-----TCTAAGATGATCTTGGAGATTATGAT 668
Db 109 ValGluAlaLeuAsnAlaValHisGlyPheIleAlaGluLysIleArgGluMetProGln 128
QY 669 AAAGAGGCTAAG----- 680
Db 129 AsnValAlaLysThrGluProValSerIleLeuGlnProGlnThrThrValAsnProAsp 148
QY 681 -----GACACAAAACGGCTGACGAGTTCCC 707
Db 149 ArgIleLysGlnThrLeuProSerSerProThrThrThrLysSerSerProSerAspPro 168
QY 708 CTG-----AAGATCTCGGCCCATATAACTTTTGTAGGGCGT 743
Db 169 MetThrThrSerArgAlaAsnGlnValLysIleIleValProAsnSerThrAlaGlyLeu 188
QY 744 CTCATTGGCAGGAAGGACCGAACCTGAGAGAGGTAGAGCAAGATACCGAGACAAAAATC 803
Db 189 IleIleGlyLysGlyGlyAlaThrValLysAlaValMetGluGlnSerGlyAlaTrpVal 208
QY 804 ACCATCTCTGTTCCAGACCTTACCTTTACACCTCTGAGAGGACCATCTACTGTGAAG 863
Db 209 GlnLeuSerGlnLysProAsp---GlyIleAsnLeuGlnGluArgValValThrValSer 227
QY 864 GGGGCCATCGAGAAATTGTTGAGGGCCGAGCAGGAAGAAATAATGAAGAAAGTTCGGAG--- 920
Db 228 GlyGluProGluGlnAsnArgLysAlaValGluLeuIleIleGlnLysIleGlnGluAsp 247
QY 921 -----GCCTATGAGAAT----- 932
Db 248 ProGlnSerGlySerCysLeuAsnIleSerTyrAlaAsnValThrGlyProValAlaAsn 267
QY 933 -----GATGTGCTGCCATG 947
Db 268 SerAsnProThrGlySerProTyrAlaAsnThrAlaGluValLeuProThrAlaAla 287
QY 948 AGCTCTCACTGATCCCTGGCCCTGAACTGGCTGTGTAGGTCTTTTCCCA----- 998
Db 288 AlaAlaGlyLeuLeuGlyHisAlaAsnLeuAlaGlyValAlaAlaPheProAlaValLeu 307
QY 998 ----- 998
Db 308 SerGlyPheThrGlyAsnAspLeuValAlaIleThrSerAlaLeuAsnThrLeuAlaSer 327
QY 999 -----GCTTCATCCAGCGCA 1013
Db 328 TyrGlyTyrAsnLeuAsnThrLeuGlyLeuGlyLeuSerGlnAlaAlaThrGlyAla 347
QY 1014 GTCCCGCCGCTCCAGCAGCGTTACTGGGGGTGCTCCCTATAGTCTCTTTATGAGGCT 1073
Db 348 LeuAlaAlaAlaAlaSerAlaAsnProAlaAlaAlaAlaAlaAsnLeuLeuAlaThr 367
QY 1074 CCGAGCAGGAGATGTGTCAGGTGTTATCCCGCCAGCGAGGTGGGCGCCATCATCGGC 1133
Db 368 TyrAlaSerGlu-----AlaSerAlaSerGlySerThrAlaGly 380
QY 1134 AAGAGGGGCGCAGCATCAACAGCTCTCCCGTTTCCAGCGCTCCATCAAGATTGCA 1193
Db 381 GlyThrAla-----GlyThrPheAlaLeuGlySerLeuAlaAla 394
QY 1194 CCACCCGAAACACCTGACTCCAAAGTTCGTATGTTATCATCTGAGCCGCGCAGAGGC 1253
Db 395 -----ThrAlaAlaThrAsnGly 400
QY 1254 CAATTCAAGGCTCAGGAAAGAAATCTATGCAAACTC-----AAGGAGAGAACTTC 1304
Db 401 TyrPheGlyAlaAlaSerProLeuAlaAlaSerAlaIleLeuGlyThrGlyLysSerThr 420
```

```
Db 408 ---SerAlaProAlaGlnAsnAsnPheGlyGlyGlnGlnGlnTrpAsnProValAlaGln 426
Qy 1662 CTCCTCCGCGCAGCCTGAGAAATGAGTGGGAATCCGGG----- 1697
Db 427 IleProAlaAlaAlaGlnAsnProTyrGlnValGlyGlyTyrGlnGlnAsnSerValTyr 446
Qy 1698 -----ACACNTGGCGCGCTGTAGATCAGGTTTG----- 1727
Db 447 AlaGlnGlnThrAlaAlaProAlaAlaAlaProTyrAlaAlaAlaGlyIleValGlnPro 466
Qy 1728 -----CCCACCTTGATTGAGAAAGATGTTCCAGTGGAGAAC 1763
Db 467 GlnGlnGlnValAlaTyrGlnGlnProGlnValValGlnGlnVal---GlnProAlaThr 485
Qy 1764 CTGATCTNTCAGCCGCCAACACCCAC----- 1790
Db 486 ThrValAlaAlaThrThrThrProThrValAspProValThrGlyGluGlnAspTyrSer 505
Qy 1791 ---CAATGGCCCAACACTGNTGCCCTCGGGTG----- 1823
Db 506 AlaGlnTrpMetGluTyrTyrLysSerIleGlyAlaHisAspLysAlaGluAlaValGlu 525
Qy 1824 -----TCAGAAATTNTAGCGCAAGGCACCTTTTAAACGTGGATTG 1862
Db 526 AlaGlnMetLysLysLysLysAlaGluAlaAlaAlaAlaAlaAlaValProGlyGlyLeu 545
Qy 1863 TTAAAGAAGCTCTCCAGGCCCCCAAGAGGGTGGATCACACCTCAG 1910
Db 546 -----IleGlnGlnMetProMetGlyMetAlaMetProGln 557

RESULT 11
T19217
hypothetical protein C12D8.1b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T19217
R;McMurray, A.
submitted to the EMBL Data Library, June 1996
A;Reference number: Z19092
A;Accession: T19217
A;Molecule type: DNA
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Residues: 1-611 <WIL>
A;Cross-references: EMBL:Z73969; PIDN:CAA98233.1; GSPDB:GN00023; CESP:C12D8.1b
A;Experimental source: clone C12D8
C;Genetics:
A;Gene: CESP:C12D8.1b
A;Map position: 5
A;Introns: 181/3; 340/3; 535/2

Alignment Scores:
Pred. No.: 3.41e-08 Length: 611
Score: 232.00 Matches: 108
Percent Similarity: 37.15% Conservative: 106
Best Local Similarity: 18.75% Mismatches: 232
Query Match: 6.57% Indels: 130
DB: 2 Gaps: 18

US-09-270-437d-7 (1-1946) x T19217 (1-611)

Qy 435 CCAGCCAGCAGCAGCAAGTGGACATCCCGCTT----- 467
Db 50 ProAlaLysArgProAlaAspAspThrAspLeuAsnProPheMetAspAsnGluAla 69
Qy 468 -----CGGTCTCTGGTGGCCACCAGTCATGTGTGGTGCATTATTGGCAAGGAGGG 518
Db 70 ValAsnGluLysTyrProIleProGluSerAlaValGlyIleValGlyArgGlyGly 89
Qy 519 GCCACCATCCGCAACATCAACAAACAGACCCAGTCCTCAAGATAGACGTATAGGAAGGAG 578
Db 90 SerGluIleGlnGlyIleGlnAlaLysAlaGlyCysArgValGlnMetSerProAspAla 109
```

```
Qy 579 AACCGAGTGTGAGTGAATAAAGCCATCATGTGTGACTCCACCCCTGAGGGGTGCTCTCC 638
Db 110 AspProSerSerGlyValArgMetValThrLeuGluGlySerArgSerAsnValGluThr 129
Qy 639 GCTTGTAAAGATGATCTTCGAGATTATGCATAAAGAGGCTAAGGACACCAAA----- 689
Db 130 AlalysHisLeuIleAsnGluValValAlaArgSerGlnAsnProArgProGlnTyrGly 149
Qy 690 ---ACGGGTGACGAGGTTCCCTCGAAGATCTCGCCCATATAATACTTTGTAGGGGTCTC 746
Db 150 PheProArgAlaGlnThrThrIleAspIleAlaIleProProAsnArgCysGlyLeuIle 169
Qy 747 ATTGGCAAGAGAGACGCACTGAAGAGGTAGAGCAAGATACCGACACAAAATCACC 806
Db 170 IleGlyLysSerGlyAspThrIleArgGlnLeuGlnLysSerGlyCysLysMetIle 189
Qy 807 ATCTCCTCTGTGCAAGAC---CTTACCCCTTTACAACTCGAGAGGACCATCACTGTGAAG 863
Db 190 Leu-----ValGlnAspAsnGlnSerValSerAspGlnSerLysProLeuArgIleThr 207
Qy 864 GGGCCCATCGAGAAATTGTCAGGGCCGAGCAGGAAATAATGAAGAAAGTTGGGAGGCC 923
Db 208 GlyAspProGlnLysIleGluLeuAlaLysGlnLeuValAlaGluIleLeuAsnSerGly 227
Qy 924 TATCAGAAATGATGTGGTGCATGAGCTCTCACCTGATCCCTGGCTGAACCTGGCTGCT 983
Db 228 GlyAspGlyAsnGlyGlySer-----GlyLeuGlnMetHis 240
Qy 984 GTAGTCTCTTTCCAGCTTCATCCAGCGCAGTCCCGCGCTCCAGCAGCGCTTACTGGG 1043
Db 241 AlaGly-----Gly 243
Qy 1044 GCTGCTCCTATAGTCTCTTTATGCAGGCTCCGAGCAGGAGATGTTGAGGTGTTTATC 1103
Db 244 GlyGlyGlyAlaSerAla-----ArgGlyGluValValVal 255
Qy 1104 CCCGCCAGCAGTGGGGCCATCATCGGCAAGAGGGGAGCAGACATCAACAGTCTCC 1163
Db 256 ProArgSerSerValGlyIleIleIleGlyLysGlnGlyAspThrIleLysArgLeuAla 275
Qy 1164 CGGTTTGCAGCGCTCTCATCAAGATTGACACCCGCAAAACACCTGATCCCAAGTTCGT 1223
Db 276 MetGluThrGlyThrLysIleGlnPheLysProAspAspProSerThrProGluArg 295
Qy 1224 ATGTTTATCATCTGAGCCGCCAGGCCCAATTCAGAGTCTAGGGAGAAATTCATGGC 1283
Db 296 CysAlaValIleMetGlyThrArgAspGlnIleTyrArgAlaThrGluArgIleThrGlu 315
Qy 1284 AAATCAAGGAGGAG-----AACTTCTTGGTCCCGAGGAG 1319
Db 316 LeuValLysLysSerThrMetGlnGlnGlyGlyGlyAlaGlyAlaMetVal 335
Qy 1320 GAAGTGAAGCTGGAGACC---CACATAGTGTGCGCAGCATCAGCAGCTGGCGGTCTATT 1376
Db 336 SerAsnGluAlaSerThrPheTyrMetSerValProAlaAlaLysCysGlyLeuValIle 355
Qy 1377 GGCAAGGTGGAAAAACGGTGAACAGGTTCAGAAATTTGACGGCAGCTGAGGTGGTAGTA 1436
Db 356 GlyLysGlyGlyGluThrIleLysGlnIleAsnSerGluSerGlyAlaHisCysGluLeu 375
Qy 1437 CCAGAGACAGACCCCTGTATGAGAACGACGAGTCTATCGTG-----AAA 1481
Db 376 SerArgAspProThrGlyAsnAlaAspGluLysValPheValIleLysGlyGlyLysArg 395
Qy 1482 ATCATCCGACATTTCTATGCGCAGTCAGATGGCTCAACGGAAGATCCGAGACATCTGGCC 1541
Db 396 AlaIleGluHis-----AlaLysHisLeuIleArgIleLysValGlyAspIleAlaPro 413
Qy 1542 CAGGTTTAAGCAGCAGCATCAGAGGGGACAGAGTAACCCAGCCCGCAGCGGAGGAGTGA 1601
Db 414 AsnThrProPheArgAspAspSerAlaMetThrMetGlnThrGlnPhe----- 429
Qy 1602 CCAGCCCTCTCTCTCTTGAGTCCAGGACACCAACGGGACAAATCGAGAGTGTGCT 1661
```

```

Db      422 AlaSerGlyTyrGlyGlnAlaGlyGlnGlyTyrGlnGln-----AspGlyTyrGlyAla 439
QY      1349 GCCAGCATCAGCAGCTGGCCGGGTCTATTGGCAAGAGGTGGAAGAAACGGTGAACGAGTTGCA 1408
Db      440 TyrAsnAlaSerGlnGlnSerGlyTyrGlyGln-----AlaAla 452
QY      1409 GAATTTGACGGCAGCTGAGGT---GGTAGTACCAAGACACGACCCCTGATGAGAACGA 1465
Db      453 GlyTyrAspGlnGlnGlyTyrGlySerThrThrAsnPro-SerGlnGlnGluAsp-- 471
QY      1466 CCAGTCTACGTGAAATCATCGGACATCTCATGCCAGTCTCAGATGGCTCAACGGAAGAT 1525
Db      472 -----AlaSerGlnAlaAlaPro----- 477
QY      1526 CCGAGACATCTCGCCAGGTTTAAGCAGCAGCATCAGAAGGGACAGAGTAACACGAGCCCA 1585
Db      478 -----ProSerSerAlaGlnSerGlyGlnAlaGlyTyrGlyTh 490
QY      1586 GGCACGGAGGAAGTACACGACCCCTCTGTCCTTNGAGTCCAGGACACACACCGGGCAG 1645
Db      490 rThrGlyGlnGlnProProAlaGlnGlySerThrGlyGlnAlaGlyTyrGlyAlaProPr 510
QY      1646 AAATCAGAGTGTGCTCTCCCGGCGAGCTCAGAAATGAGTGGAAATCCGGGACACNTGG 1705
Db      510 oThrSerGlnAlaGlyTyrSerSerGlnProAlaAlaAlaIarYrAsnSerGlyTyrGlyAl 530
QY      1706 GCGGGCTGTAGATCAGGTTCGCCACT 1733
Db      530 aProProAlaSerLysProProThr 539

RESULT 10
T19216
hypothetical protein C12D8.1a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T19216
R:McMurray, A.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19092
A:Accession: T19216
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-589 <MIL>
A:Cross-references: EMBL:Z73969; PIDN:CAA98232.1; GSPDB:GN00023; CESP:C12D8.1a
A:Experimental source: clone C12D8
C:Genetics:
A:Gene: CESP:C12D8.1a
A:Map position: 5
A:Introns: 7/3; 25/1; 159/3; 318/3; 513/2

Alignment Scores:
Pred. No.:      3.37e-08      Length:      589
Score:          232.00      Matches:     108
Percent Similarity: 37.15%      Conservative: 106
Best Local Similarity: 18.75%      Mismatches:  232
Query Match:      6.57%      Indels:     130
DB:              2          Gaps:         18

US-09-270-437D-7 (1-1946) x T19216 (1-589)
QY      435 CCAGCCAAGCAGCAGCAGTGGGACATCCCCCTT----- 467
Db      28 ProAlaLysArgProAlaAspAspThrAspLeuAsnProPheMetAspAsnGluAla 47
QY      468 -----CGGCTCTGGTGCCACCCAGTATGTGGTCCCATTTATGGCAAGAGGGG 518
Db      48 valAsnGluLysTyrProIleProGluSerAlaValGlyIleValIleGlyArgGlyGly 67
QY      519 GCCACCATCGCAACATCAAAACACCCAGTCCCAAGATACAGCATGTCATAGAAAGGAG 578
Db      68 SerGluIleGlnGlyIleGlnAlaLysAlaGlyCysArgValGlnMetSerProAspAla 87

```

```

QY      579 AACGCGAGTGCAGCTGAAAAAGCCATCAGTGTGCACCTCCACCCCTGAGGGCTGCTCTCC 638
Db      88 AspProSerSerGlyValArgMetValThrLeuGluGluGlySerArgSerAsnValGluThr 107
QY      639 GCTTGTAGATGATCTTTGGAGATTATGCATAAAGAGGCTAAGGACACCCAAA----- 689
Db      108 AlalysHisLeuIleAenGluValAlaAlaArgSerGlnAsnProArgProGlnTyrGly 127
QY      690 ---ACGGCTGACGAGGTTCCTCCTGAAAGATCTCTGGCCCCATAATAACTTTGTAGGGGTCTC 746
Db      128 PheProArgAlaGlnThrThrIleAspIleAlaIleProProAsnArgCysGlyLeuIle 147
QY      747 ATTGGCAGGAGGAGCGAACCTGGAAGAGTGTAGAGCAAGATACCGAGACAAAATACCC 806
Db      148 IleGlyLysSerGlyAspThrIleArgGlnLeuGlnGluLysSerGlyCysLysMetIle 167
QY      807 ATCTCTCTCTGTGCAGAC---CTTACCCCTTTTACAACTCCCTGAGAGGACCATCTCTGAAG 863
Db      168 Leu-----ValGlnAspAsnGlnSerValSerAspGlnSerLysProLeuArgIleThr 185
QY      864 GGGGCCATCGAAGATTGTTGAGGGCCGAGCAGGAGAAATAATGAAGAAAGTTTCGGAGGCC 923
Db      186 GlyAspProGlnLysIleGluLeuAlaLysGlnLeuValAlaGluIleLeuAsnSerGly 205
QY      924 TATGAGAATGATGTGGCTGCCATCAGCTCTCACCTGATCCCTGGCTGAACCTGCTGCT 983
Db      206 GlyAspGlyAsnGlyGlySer-----GlyLeuGlnMetHis 218
QY      984 GTAGTCTTTTCCAGCTTCATCCAGCAGCTCCCGCGCCTCCAGCAGGCGTTACTGGG 1043
Db      219 AlaGly-----Gly 221
QY      1044 GCTGCTCCCTATAGCTCTCTTTATGCGAGCTCCGAGCAGGAGATGGTCAGGTGTTATC 1103
Db      222 GlyGlyGlyAlaSerAla-----ArgGlyGluValValVal 233
QY      1104 CCGCGCCAGGAGTGGCGGCCATCATCGCAAGAGGGGAGCAGACATCAACAGCTCTCC 1163
Db      234 ProArgSerSerValGlyIleIleGlyLysGlnGlyAspThrIleLysArgLeuAla 253
QY      1164 CGGTTTGGCAGCGCTCCATCAAGATTGCACCCAGAACACCTGACTCCAAAGTTGCT 1223
Db      254 MetGluThrGlyThrLysIleGlnPheLysProAspAspProSerThrProGluArg 273
QY      1224 ATGGTTATCATCTACGACCGCCAGCAGGCCCCAATTCAGGCTCAGGAGATCTATGCG 1283
Db      274 CysAlaValIleMetGlyThrArgAspGlnIleTyrArgAlaThrGluArgIleThrGlu 293
QY      1284 AAACCTCAAGGAGGAG-----AACTTCTTGGTCCCAAGGAG 1319
Db      294 LeuValLysLysSerThrMetGlnGlnGlyGlyGlyAlaGlyAlaMetVal 313
QY      1320 GAAGTGAAGCTGGAGACC---CACATACGTGTGCCAGCATCAGACGCTGGCGGGTCATT 1376
Db      314 SerAsnGluAlaSerThrPheTyrMetSerValProAlaAlaLysCysGlyLeuValIle 333
QY      1377 GCGAAAGGTGGAAACCGTGAACGATTGTCAGATTTCACGCGCAGCTGAGGTGGTAGTA 1436
Db      334 GlyLysGlyGlyGluThrIleLysGlnIleAsnSerGluSerGlyAlaHisCysGluLeu 353
QY      1437 CCAAGAGACCAACCCCTGATGAGAACACGACAGGTCATCGTG-----AAA 1481
Db      354 SerArgAspProThrGlyAsnAlaAspGluLysValPheValIleLysGlyGlyLysArg 373
QY      1482 ATCATCGGCACATTTCTATGCCAGTCCAGTGTGTCAACGGAAGATCCGAGCATCTCTGGCC 1541
Db      374 AlaIleGluHis-----AlalysHisLeuIleArgIleLysValGlyAspIleAlaPro 391
QY      1542 CAGGTTAAGCAGCAGCATCAGAGGACAGAGTAACCGCCCGCAGCGCAGGAGAGTGA 1601
Db      392 AsnThrProPheArgAspAspSerAlaMetThrMetGlnThrGlnPhe----- 407
QY      1602 CCAGCCCTCTCCCTGTCCCTTNGAGTCCAGGACAAACAAACGGGAGAAATCGAGAGTGTGCT 1661

```



```
QY 681 -----GACACAAACGGCTGACAGAGTTCCTCCCTGAAGATCTCGGCGCCATTAATAC 731
Db 233 ProtleProAlaValSerAsnSerLysValThrIlelleProIleProAlaAsn 252
QY 732 TTTGTAGGCGCTCATTTGGCAAGGAGGACGACCTGACAGAGTGAAGGTAGACAGATACC 791
Db 253 LysCysGlyAlaIlelleGlyLysGlyGluGlnMetArgLysLeuArgSerTrpThr 272
QY 792 GAGACAAAATACCATCTCTCGTTGCAAGACCTTACCCTTTACAAACCTCGAGAGACC 851
Db 273 AsnCysAspPheIleLeu-----IleGlnGluAsnAsnIleAlaAspSerValLysPro 290
QY 852 ATCACTGTGAAGGGGCCATGACAGAAATTGTCAGGGCGGAGCAGGAATAATAGAGAAA 911
Db 291 LeuGlnIleThrGlyGln-----ProLysGlu 299
QY 912 GTTCGGAGGCTATGAGAAATGATGTGGCTGCCATGAGCTCTCACCTGATCCTCGGCTG 971
Db 300 ValGluHisAla-----LysAlaLeuValAlaAspIleLeuAspGlyPhe 314
QY 972 AACCTGGCTGTAGGTCTTTTCCAGCTTCATCCAGCGAGTCCCGCGCTCCCGC 1031
Db 315 AspGluCys-----ProProAla 320
QY 1032 AGCGTTACTGGGCTGCTCCCTATAGCTCTTATGTCAGGCTCCCGAGCAGAGATGCTG 1091
Db 321 GlyMetAlaGlyAsnSerProValAlaAlaMet-----SerLeu 333
QY 1092 CAGGTGTTTATCCCGCGCCAGGACGTGGCGGCATCATCGCAAGAGGGGAGCAGCATC 1151
Db 334 GlnValLysValProArgSerThrValGlyAlaIleMetGlyLeuGlnGlySerAsnIle 353
QY 1152 AAACAGCTCTCCGGTTGCGCGCCTCCATCAAGATTGCACACCGCGCAACACCTGAC 1211
Db 354 LysLysIleSerAsnGluThrGluThrLysIleGlnPheMetProAspAspProLys 373
QY 1212 TCCAAAGTTCGTATGTTATCATCATCTGACCGCCAGCGAGCCCAATTCAGGCTCAGGA 1271
Db 374 LeuMetGluArgThrLeuValValIleGlyAsnLysAsnLysValTyrValCys---Ala 392
QY 1272 AGAATCTATGGCAACTCAAGGAGGAGAACTCTTTGGTCCCAAGGAGGAGTGAAGCTG 1331
Db 393 ArgLeuLeuGlnIleValGluAlaAsnSerGluAsnAlaAsnThrProIleSerLeu 412
QY 1332 GAGACCCATACGTGTGCGACATCAGCAGCTGGCGGGGTCAATTGGCAAGAGTGAATA 1391
Db 413 ---PheTyrMetLeuIleProAlaSerLysCysGlyLeuValIleGlyArgGlyGly 431
QY 1392 ACGGTCAAGAGCTGCAGAAATTTGACGGCAGCTGAGTGGTAGTACCAAGAGAC 1445
Db 432 ThrIleArgGlnIleAsnLysGluSerGlyAlaTyrCysGluMetSerArgAsp 449

RESULT 8
T49962
Hypothetical protein F8M21.160 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C/Accession: T49962
R:Pevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le
submitted to the Protein Sequence Database, April 2000
A/Reference number: Z24493
A/Accession: T49962
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-568 <BEV>
A/Cross-references: EMBL:AL353993; GSPDB:GN00063; ATSP:F8M21.160
A/Experimental source: cultivar Columbia; BAC clone F8M21
C/Genetics:
A/Gene: ATSP:F8M21.160
A/Map position: 5
A/Introns: 200/3; 337/3; 544/3
```

```
Alignment Scores:
Pred. No.: 2.23e-08 Length: 568
Score: 234.50 Matches: 92
Percent Similarity: 42.47% Conservative: 97
Best Local Similarity: 20.67% Mismatches: 175
Query Match: 6.64% Indels: 81
DB: 2 Gaps: 16

US-09-270-437d-7 (1-1946) x T49962 (1-568)
QY 336 CCCGATGAGCAGATAGCAGGACCTGAGAAATGGCGCCGAGGGGGCTTTGGTCTCTCGG 395
Db 3 ProAspHisArgMetSerProAspHisArgAspSerHisArg-----LysArg 18
QY 396 GGTGACCCCGCCAGGGCTCACCTGTGGCAGCGGGGGCCCCCAGCCAGCAGCAG----- 449
Db 19 SerArgProGlnSerAspTyrAspAsnGlyGlySerLysArgArgTyrArgGlyAsp 38
QY 450 -----CMAGTGACATCCCTTCCGCTCTCGCTCTCGTGGCCACC 485
Db 39 AspArgAspSerLeuValIleAspArgAspAspThrValPheArgTyrLeuCysProVal 58
QY 486 CAGTATGTGGTGCCATTATTGGCAAGGAGGGGGCCACCTCCGCAACATCACAACACAG 545
Db 59 LysLysIleGlySerValIleGlyArgGlyGlyAspIleValLysGlnLeuArgAsnAsp 78
QY 546 ACCAGTCCCAAGATAGACGTGCATAGGAGGAGGACGAGGTGCAGCTGAAGAACCCATC 605
Db 79 ThrArgSerLysIleArgIle---GlyGluAlaIleProGlyCysAspGluArgValIle 97
QY 606 AGTGTGCACTCCACCCCT-----GAGGGCTGCTCTCTCC 638
Db 98 ThrIleTyrSerProSerAspGluThrAsnAlaPheGlyAspGlyGluLysValLeuSer 117
QY 639 GCTGTGAAGATGATCTTCGAGATATGATTAAGAG-----GCTAAGACACCAAAACG 692
Db 118 ProAlaGlnAspAlaLeuPheArgIleHisAspArgValAlaAspAlaArgSer 137
QY 693 GCTGAC-----GAGGTTCCCTGAAATCCTCGGCCATAATAACTTT 734
Db 138 GluAspSerProGluGlyGluLysGlnValThrAlaLysLeuLeuValProSerAspGln 157
QY 735 GTAGGCGCTCTCATTTGCAAGGAGGACGGAACCTTGAAGAAGGTAGACAGATACCGAG 794
Db 158 IleGlyCysIleLeuGlyArgGlyGlyGlnIleValGlnAsnIleArgSerGluThrGly 177
QY 795 ACAAATACCATCTCTCTGTCAGACACCTTACCTT-----TACAACCTCAGAGG 848
Db 178 AlaGlnIleArgIleValLysAspAspAsnMetProLeuCysAlaLeuAsnSerAspGlu 197
QY 849 ACCATCACTGTGAAGGGGCCATCGAGAATTGTCAGGGCCGAGCAGGAAATAATGAAG 908
Db 198 LeuIleGlnIleSerGlyGluValleuIleValLysLysAlaLeuGlnIleAlaSer 217
QY 909 AAAGTTCGGAGGCGCTATGAGAATCATGTGGCTGCCATGAGCTCTACCTGATCCTCGGC 968
Db 218 ArgLeu-----HisGluAsn-----ProSerArgSerGlnAsnLeu----- 229
QY 969 CTGAACCTGGCTGTGTAGTCTTTTCCAGAGTTCATCCAGCGCAGTCCCGCGGCTCC 1028
Db 230 -----LeuSerSerSerGlyGlyTyrProAlaGlySerLeuMetSerHisAlaGlyGly 247
QY 1029 AGCAGGTTACTGGGGCTGCTCCCTATAGCTCC----- 1061
Db 248 ProArgLeuValGlyLeuAlaProLeuMetGlySerTyrGlyArgAspAlaGlyAspTrp 267
QY 1062 -----TTTATGACAGGCTCCCGAGCAGGAG-----ATGTTGCGAG 1094
Db 268 SerArgProLeuTyrGlnProProArgAsnAspProProAlaThrGluPheIleArg 287
QY 1095 GTGTTTATCCCGCCCGCAGGAGTGGGGCCATCATCGGCAAGAGGGGAGCAGCATCAAA 1154
Db 288 LeuValSerProValGluAsnIleAlaSerValIleGlyLysGlyAlaLeuIleAsn 307
```

QY 861 AAGGGCCATCGAATTGTTGACGGCCGACGAGGAATAATGAAGAAGTTCCGGAG 920
Db 238 ThrGlyAspProTyrLysValGlnGlnAlaLysGluMetValLeuGluLeuIleArgAsp 257
QY 921 -----GCATGACAAATGATGTGGCTGCCATGAGCTCTCACCTGATCCCT 965
Db 258 GlnGlyGlyPheArgGluValArgAsnGluTyrGly----- 269
QY 966 GGCGTGAACCTGGCTGCTGTAGTCTTTTCCACGCTTCATCCAGCGAGTCCCGCGCCT 1025
Db 269 ----- 269
QY 1026 CCCAGCAGCTTACTGGGCTGCTCCCTATAGCTCCTTTATGCAAGGCTCCGACGACGAG 1085
Db 270 ---SerArgIleGlyGly-----AsnGlu 276
QY 1086 ATGCTGACGCTGTTTATCCCGCCAGCGAGTGGCGCCATCATCGCAAGAGGGCGAG 1145
Db 277 GlyIleAspValProIleProArgPheAlaValGlyLeuValIleGlyArgAsnGlyGlu 296
QY 1146 CACATCAACAGCTCTCCGGTTGCGAGCGCTCCATCAAGATTGACCAACCCGAA--- 1202
Db 297 MetIleLysLysIleGlnAsnAspAlaGlyValArgIleGlnPheLysProAspAspGly 316
QY 1203 ---ACACCTGACTCCAAAGTTGCTATGCTTATCATCTGACCGCCAGAG---GCCCAA 1256
Db 317 ThrThrProGlu-----ArgIleAlaGlnIleThrGlyProProAspArgCysGln 333
QY 1257 TTCAAGGCTCAG----- 1268
Db 334 HisAlaAlaGluIleIleThrAspLeuLeuArgSerValGlnAlaGlyAsnProGlyGly 353
QY 1269 -----CGAAGAANTTATGCAAACTCAAGGAGGAGAACTTCTTTGTTGCCAAG 1316
Db 354 ProGlyProGlyArgGlyArgGlyArgGlyGlnGlyAsnTrpAsnMetGlyProPro 373
QY 1317 GAGAAAGTGAAGCTGAGACCCACATACGTGTGCCAGCATCAGCAGTGGCCGGTCAAT 1376
Db 374 GlyGlyLeuGln---GluPheAsnPheIleValProThrGlyLysThrGlyLeuIle 392
QY 1377 GCGAAAGTGGAAAAACGGTGAAGATTGCAGAAATTGACGCGCAGCTGAGGTGGTAGTA 1436
Db 393 GlyLysGlyGlyGluThrIleLysSerIleSerGlnGlnSerGlyAlaArgIleGluLeu 412
QY 1437 CCAAGACACAGACCCCTGATGAGAACGACCAAGTGC---ATCGTGAATAATCATCGACAT 1493
Db 413 GlnArgAsnProProProAsnAlaAspProAsnMetLysLeuPheThrIleArgGlyThr 432
QY 1494 -----TTCTATGCCAGTCAGATGGCTCAACCGAAGAT-----CCGACACAT 1534
Db 433 ProGlnGlnIleAspTyrAlaArgGlnLeuIleGluGluLysIleGlyGlyProValAsn 452
QY 1535 CCT---GGCCAGGTTAAAGCAGCAGCATCAGAAGGGACAGAGTAACAGAGCCCGAGGACG 1591
Db 453 ProLeuGlyProProValProHisGlyProHisGlyVal-----ProGlyPro---HisG 470
QY 1592 GA-----GGAAGTGAACAGCCCTC-----CTGCTCCCTTNGA 1624
Db 470 LysProGlyProProGlyProGlyThrProMetGlyProTyrAsnProAlaProTyrA 490
QY 1625 GTCCAGGA----- 1632
Db 490 snProGlyProProGlyProAlaProHisGlyProProAlaProTyrAlaProGlnGlyT 510
QY 1633 -----CAACAACGGGCAGAAATCGAGAGTGTGCTCTCCCGCGG 1669
Db 510 xPglyAsnAlaTyrProHisTrpGlnGlnGlnAlaProProAsp-----ProA 526
QY 1670 CAGGCTGGAATGATGGGAATCCGGGACACNTGGCGGGCTGTAGATCAGGTTGCC 1729
Db 526 IalysAlaGlyThrAspProAsnSerAla-AlaTrpAlaAlaTyr-----TyrAla 542
QY 1730 CACTTGATTGAGAAAGATGTTCCAGTGAGGAACCTTGATCTNTCAGCCCCCAACACCCAC 1789

Db 543 HisTyrTyrGlnGlnGlnAla-----GlnProProProAlaAla 555
QY 1790 CCAATTGGC---CCAACACTGTTNTGCCCTCGGGTGTCAAAATTTTAGCGCAAGGCAC 1846
Db 556 ProAlaGlyAlaProThrThrThrGlnThrAsnGlyGlnGlyAsp----- 570
QY 1847 TTTTAAACGTGATGTTTAAAGAAAGCTCTCCAGGCCCCCAAGAGAGGTGTGATCACACC 1906
Db 571 -----GlnGlnAsnProAlaProAlaGlyGlnValAspTyrThr 583
QY 1907 -----TCAGTGGGAAGAAAAATAAATTTCTCTTCAGGT 1939
Db 584 LysAlaTrpGluGluTyrTyrLysLysMetGlyGlnAlaValProAlaProThrGly 602
RESULT 7
T25832
hypothetical protein M01A10.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T25832
R;Sheet, P.
submitted to the EMBL Data Library, February 1997
A;Description: The sequence of C. elegans cosmid M01A10.
A;Reference number: Z20094
A;Accession: T25832
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-680 <SCH>
A;Cross-references: EMBL:U88174; PIDN: AAB42272.1; GSPDB: GN00019; CESP: M01A10.1
A;Experimental source: strain Bristol N2; clone M01A10
C;Genetics:
A;Gene: CESP: M01A10.1
A;Map position: 1
A;Introns: 63/2; 97/3; 121/1; 160/1; 269/3; 411/3; 512/1; 649/2
Alignment Scores:
Pred. No.: 8,98e-09 Length: 680
Score: 240.50 Matches: 82
Percent Similarity: 41.21% Conservative: 82
Best Local Similarity: 20.60% Mismatches: 159
Query Match: 6.81% Indels: 75
DB: 2 Gaps: 12
US-09-270-437D-7 (1-1946) x T25832 (1-680)
QY 342 GACAGATGACACAGGACCTGAGATGGCGCGAGGGGCTTGGCTCTCGGGTCAG 401
Db 97 LysHisIleSerSerSerThrAspThrAlaArgLysArgAspPheAspGluArgSer--- 115
QY 402 CCGCGCAGGCGCTCACCT-----GTGGCAGCGGGGGCCCCAGCAAGCAGCAAGTG 455
Db 116 -----GluGlySerAspGluTyrGluGluTyrAlaProProCysLysLeuThrLysGly 133
QY 456 GACATCCCCCTTCGG-----CTCCTGGTGGCC 482
Db 134 AspileAspTyrArgValAspThrSerThrThrValIleLysAlaSerValSerIlePro 153
QY 483 ACCAGTATGTGGTGCCATTATTGGCAAGGGGGGCCACCATCCGCAACATCACAAAA 542
Db 154 GluGluSerValGlyLeuValIleGlyArgAsnGlyValGluIleGlnAlaIleSerGln 173
QY 543 CAGACCCAGTCCCAAGATAGACGTGCATAGGAAGAGAAACGAGGTGCAGCTGAAAAAGCC 602
Db 174 LysSerGlyCysArgValGlnIle---ValAlaGluProSerThrThrGlyTyrArgSer 192
QY 603 ATCAGTGTGACATCCACCCCTGAGGGCTGCTCGCTTGTGAAGATGATCTTGGAGATT 662
Db 193 ValAspileTyrGlyLysSerGluAsnIleGluValAlaLysLysLeuIleAsnGluVal 212
QY 663 ATCATAAAGAGCTTAAG----- 680
Db 213 ValAlaArgGlyArgLysLeuSerGlnGluProLeuProCysSerValProGlnPheGln 232

102	Hi	Phe	Gly	Asp	Asp	Gly	Glu	Leu	Val	Cys	Pro	Ala	Leu	Asp	Ala	Leu	Phe	Lys	Val	His	121	
657	G	A	G	A	T	T	A	T	A	G	A	A	A	G	G	G	T	---	A	A	G	701
122	Asp	Met	Val	Val	Ala	Ala	Asp	Ala	Asp	Gln	Asp	Asp	Gly	Thr	Asp	Asp	Asn	Asp	Leu	Gly	141	
702	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	749	
142	Gl	Lys	Gln	Thr	Val	Thr	Val	Arg	Met	Leu	Val	Pro	Ser	Ser	Asp	Gln	Ile	Gly	Cys	Val	161	
750	G	C	A	A	G	A	A	G	G	C	G	A	C	T	T	G	A	A	A	A	A	809
162	Gly	Lys	Gly	Gln	Val	Ile	Gln	Asn	Leu	Leu	Arg	Asn	Asp	Thr	Asn	Ala	Gln	Ile	Arg	Val	181	
810	T	C	T	C	G	T	T	G	C	A	A	G	---	---	---	---	---	---	---	---	851	
182	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	199	
852	A	T	C	A	C	T	G	T	G	A	A	G	G	G	C	A	T	G	A	A	911	
200	L	e	u	t	i	e	l	e	g	l	y	P	r	o	L	e	u	---	---	---	207	
912	G	T	T	C	G	G	A	G	G	C	---	---	---	---	---	---	---	---	---	---	953	
208	V	al	A	r	g	A	l	a	L	e	u	T	h	r	G	l	n	V	al	A	227	
954	C	A	C	T	G	A	T	C	C	T	G	G	C	T	G	A	A	C	T	T	1013	
228	H	i	s	L	e	u	---	---	---	---	---	---	---	---	---	---	---	---	---	---	237	
1014	G	T	C	C	G	G	C	T	C	C	A	G	A	G	G	T	T	A	C	T	1070	
238	M	e	t	H	i	s	G	l	n	P	r	o	G	l	y	A	l	a	M	e	257	
1071	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1109	
258	A	l	a	V	a	l	A	r	g	A	s	p	I	l	e	A	a	S	p	A	277	
1110	C	A	G	C	A	G	G	C	G	C	A	T	C	A	T	C	G	G	A	A	1169	
278	G	l	u	a	n	V	a	l	G	l	y	V	a	l	I	e	G	l	y	S	297	
1170	G	C	A	G	C	G	C	T	C	A	A	G	A	T	T	G	C	A	C	C	1229	
298	T	h	r	G	l	y	A	l	a	T	h	r	I	e	Arg	V	a	l	Asn	Thr	315	
1230	A	T	C	A	C	T	G	A	C	C	C	C	A	G	---	---	---	---	---	---	1259	
316	I	l	e	P	h	e	l	e	S	e	r	S	e	r	L	y	s	G	l	u	335	
1260	A	A	G	C	T	A	G	G	A	A	G	A	A	T	C	T	A	T	G	G	1316	
336	A	r	g	L	e	u	G	l	n	A	r	g	C	y	s	e	r	G	l	u	350	
1317	G	A	G	A	G	T	G	A	C	T	G	G	A	C	C	A	T	A	C	G	1376	
351	---	Asp	Leu	Ala	I	e	S	e	r	Thr	Arg	Leu	Leu	Val	Ser	Ser	Ser	G	l	369		
1377	G	C	A	A	G	T	G	A	A	A	C	G	T	G	A	A	T	T	G	A	1433	
370	Gly	Lys	Gly	Gly	Ala	Val	Ile	Ser	Glu	Met	Arg	Ser	Val	Thr	Arg	Ala	Asn	Ile	Arg	389		
1434	G	T	A	C	A	A	G	A	C	C	A	C	C	T	G	A						

```

RESULT 6
A53184
myc far upstream element-binding protein - human
N:Alternate names: FUSE-binding protein
C:Species: Homo sapiens (man)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
E:Accession: A53184
R:Duncan, R.; Bazar, L.; Michelotti, G.; Tomonaga, T.; Krutzsch, H.; Avigan
Genes Dev. 8, 465-480, 1994
A:Title: A sequence-specific, single-strand binding protein activates the f
A:Reference number: A53184; MUID:94170991; PMID:8125259
A:Accession: A53184
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-644 <DUN>
A:Cross-references: GB:U05040; NID:9460151; PIDN:AAA17976.1; PID:g460152
C:Keywords: DNA binding

```

Alignment Scores:		
Pred. No.:	5.04e-09	644
Score:	244.00	148
Percent Similarity:	36.46%	93
Best Local Similarity:	22.39%	219
Query Match:	6.91%	202
DB:	2	33
US-09-270-A3D7-D (1-1946)	x A53184 (1-644)	

QY	309	AACCATGCCCTGAAGGTCCTCTACATCCCGCATGACGATAGACAG-----GGACCT	362
Db	28	AsnAspAlaPheLysAspAlaLeuGlnArgAlaArgGlnIleAlaAlaLysIleGlyGly	47
QY	363	GAGAAATGG-----CGCCGAGGGGCTTGCTCTCGGGGTTCAGCCCGCCAGGCG	413
Db	48	AspAlaGlyThrSerLeuAsnSerAsnAspTyrGlyTyrGlyGlnLysArg-----	65
QY	414	TCACCTGTGGCAGCGGG-----	431
Db	66	---ProLeuGluAspGlyAspGlnProAspAlaLysLysValalaProGlnAsnAspSer	84
QY	432	-----GCCCCAGCCCAAGCAGCAGCAAGTGGACATCCCCCTT-----CGG	470
Db	85	PheGlyThrGlnLeuProProMetHisGlnGlnSerArgSerValMetThrGluGlu	104
QY	471	CTCCTGTGTCCACCACCATGATCTGGGTGCATATTATGGCAAGAGGGGGCCACCATCCGC	530
Db	105	TyrLysValProAspGlyMetValGlyPheIleIleGlyArgGlyGluGlnIleSer	124
QY	531	AACATCACAAACAGACCCAGTCCCAAGATAGACGTGCATAGGAGGAGCAACGCGATGCA	590
Db	125	ArgIleGlnGlnGlnSerGlyCysIleGlnIle---AlaProAspSerGlyGlyLeu	143
QY	591	GCTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCTCCGCTGTAAGATG	650
Db	144	ProGluArgSerCysMetLeuThrGlyThrProGluSerValGlnSerAlaLysArgLeu	163
QY	651	ATCTTTGGATTATGCAATAA-----GAGGCTAAGGACCAACCAACGGCTGAC-----	698
Db	164	LeuAspGlnIleValGlnLysGlyArgProAlaProGlyPheHisGlyAspGlyPro	183
QY	699	---GAGGTTCCTGAAGATCTCGGCCATAATAACTTTGTAGGGCGTCTCATTTGCCAAG	755
Db	184	GlyAsnAlaValGlnGlnIleMetIleProAlaSerLysAlaGlyLeuValIleGlyLys	203
QY	756	GAAGCAGCGAACCTGAAGAAGGTAGCAGATACCGAGACAAAATACCATTCTCCTCG	815
Db	204	GlyGlyGluThrIleLysGlnLeuGlnGluArgAlaGlyValLysMetValMet-----	221
QY	816	TTGCAAGACCTTACCTTTTACACCTGAGAGGACC-----ATCAGTGTG	860
Db	222	IleGlnAsp-----GlyProGlnAsnThrGlyAlaAspLysProLeuArgIle	237


```
QY 438 -----GCCACAGCAGCAGCAAGTGGACATCCCTTGGCTCTGGTCCGCCACCCAGTAT 491
Db 88 ThrTyAlaThrGlnGln-----LeuThrLeuArgAlaLeuLeuSerThrArgGlu 104
QY 492 GTGGGTGCATTATTGGCAAGGAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAG 551
Db 105 AlaGlyIleIleGlyAlaGlyLysAsnValAlaGluLeuArgSerThrThrAsn 124
QY 552 TCCAAATAGACGTGTAGTAGAAGAGAACGCGAGGTGCAGCTGCAAAAACCCATCAGTGTG 611
Db 125 ValLysAlaGlyVal---ThrLysAlaValProAsnValHisAspArgValLeuThrIle 143
QY 612 CACTCACCCCTGAGGGCTGCTCCCTCCGCTTGAAGATCATCTTGAGATTATGATATAA 671
Db 144 SerGlyProLeuGluAsnValValAlaArgAlaTyArgPheIleAspIlePhe----- 161
QY 672 GAGGCTAAGCAGCACAAACGGCTGACGAGGTTCCTC----- 707
Db 162 ---AlaLysAsnSerThrAsnProAspGlyThrProSerAspAlaAsnThrProArgLys 180
QY 708 CTGAAGATCCTGGCCCATTAATCTTGTAGGGCGTCTCATTTGGCAAGGAGGACCGAAC 767
Db 181 LeuArgLeuLeuIleAlaHisSerLeuMetGlySerIleIleGlyArgAsnGlyLeuArg 200
QY 768 CTGAAGAAGTAGACAGATACCGAGACAAATAATCACCATCTCTCTGTTGCAAGACCTT 827
Db 201 IleLysLeuIleGlnAspLysCysSerCysArgMetIleAlaSer-----LysAspMet 218
QY 828 ACCCTTTACACCTCAGAGGACCATCACTGTGAAGGGGCCATCAGAAATGTTGCAGG 887
Db 219 LeuProGlnSerThrGluArgThrValGluIleHisGlyThrValAspAsnLeuHisAla 238
QY 888 GCCGAGCAGGAATATATGAAGAAGTTCGGAGGCGCTATGAG----- 929
Db 239 AlaIleTIPGluIleGlyLysCysLeuIleAspAspTipGluArgGlyAlaGlyThrVal 258
QY 930 -----AATGATGTGCTCCATGAGTCTCACCTGATCCCTGGCTGAACCTGCTGCT 983
Db 259 PheTyAsnProValSerArgLeuThrGln-----ProLeuProSerLeuAlaSer 275
QY 984 GTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCGCT-----CCCAGCAGCGTT 1037
Db 276 Thr-----AlaSerProGlnGlnValSerProProAlaAlaProSerThrThr 291
QY 1038 ACTGGGCTGCT-----CCCTATAGTCTC----- 1061
Db 292 SerGlyGluAlaIleProGluAsnPheValSerTyArgAlaGlnValPheProAlaThr 311
QY 1062 -----TTATGAGGCTCCGAGCAGGAGATGGTGCAGGTGTTATCCCGCCAG 1112
Db 312 GlnMetProPheLeuGlnGlnProLysValThrGln---AsnIleSerIleProAlaAsp 330
QY 1113 GCAGTGGGCGCCATCATCGCAAGAGGGGCGAGCACATCAACAGCTCTCCCGTTTGGC 1172
Db 331 MetValGlyCysIleIleGlyArgGlySerLysIleSerGluIleArgThrSer 350
QY 1173 AGGCGCTCATCAAGATTGACACACCGCAACACCTGACTCCAAAGTTTGTATGGTTATC 1232
Db 351 GlySerLysIleSerIleAlaLysGluProHisAspGluThrGlyGluArgMetPheThr 370
QY 1233 ATCACTGGACCGCCAGGCGCCCAATTCAGGCTCAGGAGAGATCTATGCAACCTCAAG 1292
Db 371 IleThrGlyThrHisGluGluAsnGluLysAlaLeuPheLeuLeuTyArgGlnLeuGlu 390
QY 1293 GAGGAG 1298
Db 391 MetGlu 392
```

RESULT 4

T27855
hypotheical protein ZK418.9 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27855

R:Fulton, L.

A:Description: The sequence of C. elegans cosmid ZK418.

A:Reference number: Z20430

A:Accession: T27855

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-768 <FUL>

A:Cross-references: EMBL:U00047; PIDN:AAA50693.1; CESP:ZK418.9

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:ZK418.9

A:Introns: 166/1; 221/3; 243/1; 279/1; 387/3; 490/1; 567/3; 727/2

Alignment Scores:

Pred. No.:	9,06e-10	Length:	768
Score:	255.00	Matches:	154
Percent Similarity:	37.13%	Conservative:	110
Best Local Similarity:	21.66%	Mismatches:	292
Query Match:	7.22%	Indels:	155
DB:	2	Gaps:	26

US-09-270-437D-7 (1-1946) x T27855 (1-768)

QY 23 GTGCTCTGTCCTCCCTTCCCTTGGCGCTGGCGCTCAGCCACCCAGAGCCCGGGTGGGA 82

Db 95 AlaIlePheSerThrSerMetArgValPheAsnGluLysLysSerTyMetHisPhePro 114

QY 83 GGGCGAGTGTCTCAGCTTCCCGGGTTAGGACGGGAAATTCAAATCCGAAATATTCCACC 142

Db 115 GlyAsnValAspIlePheProGluHisSerSerGlnLeuPheAsnProHisPheLeuAsn 134

QY 143 CCAGTCCGATGGGAGTACTGACAGCGCTGCTGCTCAGTATGTTACAGTAGAAC-- 200

Db 135 -LysIleArgSerGlnLeuIleGlyLysLeuIle---GluLeuAspIleThrLysAsnG1 153

QY 201 -----TGTCAGCAAGTGAACACCGAGTGAACCGC 232

Db 153 uSerLeuLeuLeuSerAlaLeuPheLeuCysSerAlaValHisProAspIleSerThrG1 173

QY 233 AGTGTGTGATGTCCTCTATTCACCGGGAGCAGACCCAGGCAAGCATCATG-- 284

Db 173 uGlyLysAsnLeuLeuTyLysTyGlnGlnTyThrSerAlaAlaLeuLeuAsnHisCy 193

QY 285 -----AAGCTGAATGGCCACCGAGTGGAGAACCATGCCCTGAAGTCTCTTA 331

Db 193 sCysLeuThrAsnGlnAsnAlaProSerArgTySerGluLeuLeuSerLeuTyG1 213

QY 332 CATCCCGATGACAGATAGACAGGACCTGAGATGGCGCCGCGGGGGCTTTGGCTC 391

Db 213 nIleLeuGluGlu-----ThrHisGlnLysIleIleGlyAlaAspGlyGlyIleGlyTh 231

QY 392 TCGGGGTGAGCCCGCCAGGGCTCCTGTGGCAGCGGG-----GCCCGACCCAGCA 445

Db 231 rLeuLysArgProLeuAsp---SerGluLeuLeuAspGlyAspLeuLeuProThrLysIy 250

QY 446 GCAGCAA-----GTGGCAATCCCTCT 466

Db 250 sSerSerGluValGlyAspLeuAsnMetGlyAspSerAspLysIleThrAspIle----- 268

QY 467 TCGGCTCTGGTGGCCACCCAGTATGTTGCCATTATGGCAAGGAGGGGGCCACCAT 526

Db 269 ----TyrProValProGluLysValGlyLeuValIleGlyLysGlyGlySerGluI1 287

QY 527 CCGCAACATCACAAACAGACCCAGTCCAAAGATAGCGTAGCAGAGAGAGACGAGC 586

Db 287 eArgLeuIleGlnGlnThrSerGlyCysArgValGlnMetAspProAspHisGlnSerVa 307

QY 587 TGCAGCTGAAAAAGCCATCAGTGTGCATCTCCACCCCTGAGGGCTGCTCTCGCTTGTA 646

Db 307 lAsnGlyPheArgAsnCysThrIleGluGlyProProAspGlnValAlaValAlaArgG1 327

C;Genetics:

A;Gene: F19C24.19

A;Map position: 1

Alignment Scores:

Pred. No.:	1,71e-10	Length:	621
Score:	265.00	Matches:	117
Percent Similarity:	38.86%	Conservative:	74
Best Local Similarity:	23.68%	Mismatches:	184
Query Match:	7.50%	Indels:	119
DB:	2	Gaps:	19

US-09-270-437D-7 (1-1946) x D96554 (1-621)

```

Qy 459 ATCCCCCTCGCTCTGCTGTCGCCACCCAGTATGTGGTGCCATTATGGCAAGAGGGG 518
Db 19 ValHisPheArgLeuLeuCysProAlaThrArgThrGlyAlaIleIleGlyLysGlyGly 38
Qy 519 GCCACCATCCGCACATCACAAAACAGACAGCCAGTCCAGATA----- 560
Db 39 SerValIleArgHisLeuGlnSerValThrGlySerLysIleArgValIleAspAspIle 58
Qy 560 ----- 560
Db 59 ProValProSerGluGluArgValValLeuIleIleAlaProSerGlyLysLysAsp 78
Qy 561 -----GACGTGCATAGAGAGAGAGACGCGAGTGCAGCTGAA----- 596
Db 79 GluSerAsnValCysAspSerGluAsnProGlySerGluGluProLysGlnGluLysGly 98
Qy 597 -----AAAGCCATCAGTGTGCACTCCACCCTGAGGCTCTCTCCGCTTGTAA 647
Db 99 SerGluCysAlaGlyThrSerGlyGlyAspAspGluGluAlaProSerSerAlaGlnMet 118
Qy 648 ATGATCTTGAGATATATGCAATAA-----GAGGCTAAGGACACCAAAACGGCTGAC--- 698
Db 119 AlaLeuLeuArgValPheGluArgIleValPheGlyAspAspAlaAlaThrValAspGly 138
Qy 699 -----GAGGTCCCTGAGATCTCTGGGCCCATATAAATCTT 734
Db 139 AspGluLeuAspLysGlyGluSerGluGlyLeuCysArgMetIleValArgGlyAsnGln 158
Qy 735 GTAGGGGTCTCATTGTCAGAGGAGGACGACGACCTGAGAGGTAGAGCAATACCGAG 794
Db 159 ValAspTyrLeuMetSerLysGlyGlyLysMetIleGlnLysIleArgGluAspSerGly 178
Qy 795 ACAAAATACCATCTCTCGTTGCAA-----GACCTT 827
Db 179 AlaIleValArgIleSerThrAspGlnIleProProCysAlaPheProGlyAspVal 198
Qy 828 ACCCTTTACACCTGAGAGGACCATCATCTGTGAAGGGGGCC-----ATCGAGAAT 878
Db 199 ValIleGlnMetAsnGlyLysPheSerSerValLysLysAlaLeuLeuValThrAsn 218
Qy 879 TGT-----TGACGGGCGGACAGGAATA 902
Db 219 CysLeuGlnGluSerGlyAlaProProThrTrpAspGluCysProPheProGln----- 236
Qy 903 ATGAAGAAAGTTCGGGAGGCTATGAGATGATGCTGCATGCTCCTCACCTGATC 962
Db 237 -----ProGlyTyrProProGluTyrHisSerMetGluTyrHis----- 249
Qy 963 CTGGGCTGAACCTGGCTGTGTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG 1022
Db 250 -----ProGlnTrpAspHis-----ProPro 256
Qy 1023 CTTCCACAGCGTGTACTGGGGCTGCTCCCTATAGCTCCTTTATGCGAGGCTCCGAGCAG 1082
Db 257 ProAsnProMetProGluAspValGlyProPheAsnArgProVal-----ValGluGluGlu 275
Qy 1083 GAGATGTGTAGGTGTATTATCCCGCGCGGACGTCAGTGGCGCATCATCGGCAAGAGGGG 1142
Db 276 ValAlaPheArgLeuLeuCysProAlaAspLysValGlySerLeuIleGlyLysGlyGly 295

```

```

Qy 1143 CAGCACATCAAAACAGCTCTCCCGGTTTCCAGCGGCTCCATCAAGATTGCACCCGAA 1202
Db 296 AlaValValArgAlaLeuGlnAsnGluSerGlyAlaSerIleLysValSerAspPro--- 314
Qy 1203 ACACCTGACTCCAAAGTTTGGTATGTTATCATCTACTGGACCGCCAGAGCCCAATTCAAG 1262
Db 315 ThrHisAspSerGluGluArgIleValIleSerAlaArgGluAsnLeuGluArgArg 334
Qy 1263 -----GCTCAGGGA-----AGAACTCTATGGCAAACTCAAGAGGAGAGAACTTC 1304
Db 335 HisSerLeuAlaGlnAspGlyValMetArgValHisAsnArgIleValGluIleGlyPhe 354
Qy 1305 TTTGGTCCCAAGAGGAGTGAAGCTGGAGAGCCACATACATGCTGCGCCAGATCAGCAGCT 1364
Db 355 -----GluProSerAlaAlaValAlaAlaArgLeuLeuValHisSerProTyrIle 371
Qy 1365 GCGCGGTCTATGGCAAGGTGGAAAAACGCTGAACGATGTGCAGATTTCCAGCGCAGCT 1424
Db 372 GlyArgLeuLeuGlyLysGlyHisLeuIleSerGluMetArgAlaThrGlyAla 391
Qy 1425 GAGGTG---GTAGTACCAAGAGACGACACCCCTGATGAGAACGACCCAG-----GTCA 1475
Db 392 SerIleArgValPheAlaLysAspGlnAlaThrLysTyrGluSerGlnHisAspGluIle 411
Qy 1476 GTGAAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATC 1535
Db 412 ValGlnValIleGlyAsnLeu-----LysThrValGlnAspAlaLeuPheGlnIle 428
Qy 1536 CTGSCCCAGGTTAAGCAGCAGCATCAAGAGGACAGATACCAAGGCCAGCGCCAGCGAGG 1595
Db 429 LeuCysArgLeuArgGluAlaMetPheProGly-ArgLeuProPheGlnGlyMetGlyGly 448
Qy 1596 AAGTGACGACGCCCCCTC-----CCTGCTCCCTTNGAGTCCAGGACA 1634
Db 448 YProProProPheMetGlyProTyrProGluProProProPheGlyProArgGly 468
Qy 1635 ACAACGGGCGAAATCGAGAGTGTCTCTCCCGCGCAGGC 1674
Db 468 nTyrProAlaSerProAspArgTyrHisSerProValGly 481

```

RESULT 3

T41600

probable pre-mRNA splicing factor - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C;Accession: T41600

R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Boche, G.; Pohl, T.

submitted to the EMBL Data Library, October 1998

A;Reference number: Z22003

A;Accession: T41600

A;Status: preliminary; translated from GR/EMBL/DBD

A;Molecule type: DNA

A;Residues: 1-398 <LVN>

A;Cross-references: EMBL:AL031825; PIDN:CAA21234.1; GSPDB:GN00068; SPDB:SPCC757.09C

A;Experimental source: strain 972h-; cosmid c757

C;Genetics:

A;Gene: SPDB:SPCC757.09C

A;Map position: 3

Alignment Scores:

Pred. No.:	5,89e-10	Length:	398
Score:	256.50	Matches:	93
Percent Similarity:	44.74%	Conservative:	60
Best Local Similarity:	27.19%	Mismatches:	130
Query Match:	7.26%	Indels:	59
DB:	2	Gaps:	14

US-09-270-437D-7 (1-1946) x T41600 (1-398)

Qy 399 CAGCCCGGC-----CAGGGCTCACCTGTGTGGACGGGGGCCCA----- 437

```

Db 68 GlnProGluProThrSerGlnValProPheSerAlaLysProProMetAspAspAla 87

```

```
QY 371 GCGCCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGCCAGGGCTCACCCTGTGGCGAGCGG 430
Db 297 -----IleProGlnAsnLeuMe 302
QY 431 GSCCCGACGACGACAGTGGACATCCCTTCGGCTCCTGGTCCCGCCACCCAGTA 490
Db 302 tMetPro---ArgCysMetLeuLysAspTrpProIleArgCysValValGluGlyLys 321
QY 491 TGTGGGTGCATTATTGGCAAGAGGGGCCACCATCCGCAACATCACAAAAACAGACCCA 550
Db 321 rHisAlaValIleGlyProAsnGlySerThrIleLysAspIleAlaSerSerThrAr 341
QY 551 GTCCAGATAGAC-----GTGCATAGGAGGAGAACGCA-----GGTGCAGCTCA 595
Db 341 gCysArgValAspPheValAsnLeuSerLysGluArgThrValLeuLysGlyAsnAs 361
QY 596 AAAAGCCATCAGTGTGCATCTCCACCCCTCAGGGCTGCTCCTCGCTGTGAAGATGATCTT 655
Db 361 pArgIleLeuThrValHisGlyValAlaGluGlnAlaThrLysAlaValAlaArgIle 381
QY 656 GGAGATTATGCATAAGAGGCT-----AAGGACACCAAAACGGCTGACGAGGTCCCT 709
Db 381 uAspValIleGlnSerGluAlaValLysAspValAsnValGlyAlaAspThrVal 401
QY 710 GAAGATCTCGGCCATAATACTTTGTAGCGGTCTCATTTGGCAGGAAGGACGGAACCT 769
Db 401 uArgMetArgAlaHisAsnGlnLeuCysGlyArgLeuIleGlyLysAlaGlySer 421
QY 770 GAAGAGGTGAGACGATATCCGACAGAAAAATCACCATCTCC-----812
Db 421 eLysGluIleMetGlnLysThrGlyThrAsnIleThrValThrLysTyrlleGluPro 441
QY 813 ----TCGTTGCAAGACCTTACC-----CTTTCAACCCCTGAGAGGACCATCAC 856
Db 441 oGlyGlyIleSerGlyLeuThrAlaAsnGluLeuLysGlyLeuMetGluArgThrIle 461
QY 857 TGTGAAGGG---GCCATCAGAAATTGTTGCAGGCGCGACGACGGAATATGAAGAACT 913
Db 461 tValArgGlyProSerIleGluAlaValValGlnAlaGluAlaLeuIleSerAla 481
QY 914 TCGGAGGCGCTATGAGAATCATGTGCTGCATGAGCTCTCACCTG-----ATCCC 964
Db 481 uLysLysCysTyrgluSerAspSerGlnLeuArgAlaGlnSerMetGlnCysProMet 501
QY 965 TGGCTGAACCTGGCTGTGTAGTCTTTCCAGCTTCATCCAGCGCGAGTCCCG-----1019
Db 501 oProMetMetProProIleLeuProGlyAlaSerSerAlaValSerAlaPr 521
QY 1020 -----CGCCTCCGACGACGCTTACTGGGGTGTCTCCCTAT-----1055
Db 521 oHisPheIleProValGlyValMetGlnValGlnHisPheAlaSerSerGlnHi 541
QY 1056 -----AGTCCTTTATGCAG-----1070
Db 541 sLeuValHisGlnAsnAlaAsnAsnSerPheLeuGlnProGlyValLeuGlnIleGlnPr 561
QY 1071 ----GCTCCGACGACGAGATGTGAGTGTATTATCCCGCCCGCAGCGAGTGGCCCAT 1126
Db 561 oGlyThrThrAsnLeuArgGlnValArgMetTrpValProAspSerMetIleGlyAla 581
QY 1127 CATCGCAAGAAGGGCAGCACATCAACAGCTCTCCCGTTTGGCAGCGCCTCCATCAA 1186
Db 581 uIleGlyAlaLysGlyLysAsnIleLysMetIleArgAspThrGlyAlaSerVal 601
QY 1187 GATTGCACCCGGAACACTGACTCCAAA-----1217
Db 601 sIleGluAlaProGluGluLysThrGlnArgGluAlaGluGluAlaGluLysLysArg 621
QY 1217 -----1217
Db 621 sLeuAspGluThrAspSerGlyCysGluGlyValAlaSerGlyAspHisProGlnGlu 641
QY 1217 -----1217
```

```
Db 641 eLeuGluAspAsnAlaThrIleAsnSerSerAspAlaIleGluGluLysProLysProVa 661
QY 1218 ----GTTCTGTATGGTTATCATCTCAGTGGACCGCCAGAGGCCCAATTCAGGCTCAGGGAAG 1273
Db 661 lSerGluArgMetValThrIleAsnGlyAspAspLeuGlnLeuLysAlaGlnSerTy 681
QY 1274 AATCTATGCAAACTCAAGAGGAGAACTTCTTTGGTCCC-----1313
Db 681 rValPheSerLysIleAlaGluThrSerSerSerLeuProSerSerGlyMetAspGly 701
QY 1314 ----AAGGAGGAGGAAGTGGAGACCCACATCGTGTGCCAGCATCAGCAGCTGGCGG 1369
Db 701 pArgSerHisMetLeuArgIleArgThrGluValSerValProThrArgIleIleGly 721
QY 1370 GGTCAATTGGCAAGGTGAAAAACGGTGAACAGTTGCAGAAATTTACGGCGAGCTGAGGT 1429
Db 721 gIleIleGlyLysGlyGlyGlnAsnValArgGluLeuGlnArgIleThrGlyAlaVal 741
QY 1430 GGTAGTACCAAGACACCAACCCCTGAT-----GA 1459
Db 741 lNysIleProGluGluGluArgAsnGlyGlyGluValTyrArgHisAspAspGlyLeu 761
QY 1460 GAACCAAGGTCATCGTGAATAATCATCGGCATTTCTATGCCAGTCCAGATGGCTCAACG 1519
Db 761 uGluAspMetThrMetIleArgThrIleGlyAsnMetTyrSerThrHisAsnValGln 781
QY 1520 GAAGATCCGAGACATCTCGGCCCGAGTTAAGCAGCAGCATCAGAGGAGCAGAGTAACA 1579
Db 781 eArgLeuAlaHisLeu-----ValAsnGluTyrTyrArgSerGlyAspHisArg 798
QY 1580 GCGCCAGGACGAGGAGAGTACCAGCCCTCCCTGTCCCTINGAGTCCAGGACACAC 1639
Db 798 nLysSerSerAspTyrLys-----804
QY 1640 GGGCGAATAATCAGAGTGTGCTCTCCCGCGCAGGCTCAGAAATGAGTGGAAATCCGGGAC 1699
Db 805 -----GlyGlyArgProHisSerAlaProSerSerGlyG 816
QY 1700 ACNTGGCGCGGCTGTAGATCATCGTTGGCCCTACTGATTGAGAAAGATGTTCCAGTGAG 1759
Db 816 nGluLysAspGlySerAla-----LeuGluLysMetAspGlnLeu 830
QY 1760 AACCTGTATCTNTCAGCCCCAACCAACCCCAATGGCCCAACACTGNTGCCCTCGG 1819
Db 830 yThrIle-----AlaProLysSerAsnSerAsnArgAlaSerProLys 845
QY 1820 GGTGTCA 1826
Db 845 rValSer 847
RESULT 2
D96554
hypothetical protein F19C24.19 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D96554
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huiziar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luross, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96554
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-621 <STO>
A:Cross-references: GB:AE005173; NID:g11094762; PIDN:AAG29695.1; GSPDB:GNC0141
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 23, 2004, 10:55:36 ; Search time 29.4003 Seconds
(without alignments)
12733.810 Million cell updates/sec

Title: US-09-270-437D-7

Perfect score: 3531

Sequence: 1 gctgtacgagggctggg.....attcttcagggttttaaa 1946

Scoring table:

	BLOSUM62	Xgapop 10.0	Xgapext 0.5
Xgapop 10.0			0.5
Xgapext 0.5			
Ygapop 6.0			7.0
Delep 6.0			

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 566732

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2 1/USPTO.spool p/US09270437/runat 23072004.095740.27439/app.query.fasta_1.5582
-DB=PIR_78 -QFMT=fastan -SUFFIX=xpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORES=spect -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HBAPEXT=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09270437 @CGN 1 1 134 @runat 23072004.095740.27439 -NCPU=3
-NO MWAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
PIR_78:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	458.5	13.0	854	2 T23837	hypothetical prote
2	265	7.5	621	2 D96554	hypothetical prote
3	256.5	7.3	398	2 T41600	probable pre-mRNA
4	255	7.2	768	2 T27855	hypothetical prote
5	251	7.1	479	2 C86275	hypothetical prote
6	244	6.9	644	2 A53184	myc far upstream e
7	240.5	6.8	680	2 T25832	hypothetical prote
8	234.5	6.6	568	2 T49962	hypothetical prote
9	232.5	6.6	632	2 T02627	hypothetical prote
10	232	6.6	589	2 T19216	hypothetical prote
11	232	6.6	611	2 T19217	hypothetical prote
12	221.5	6.3	510	2 T38489	onconeural ventral
13	218.5	6.2	356	2 S58529	alpha-complex prot
14	218	6.2	1268	2 A44125	high density lipop

15	217.5	6.2	396	2 S41224	hnRNP protein - Af
16	214	6.1	313	2 T48439	probable RNA-bind
17	209	5.9	397	2 T30168	hypothetical prote
18	205.5	5.8	846	2 T04533	hypothetical prote
19	202.5	5.7	1270	2 S23464	vigilin - chicken
20	200.5	5.7	1279	2 T41389	rna binding protei
21	199	5.6	365	2 S42471	hnRNP protein E2 -
22	197.5	5.6	463	2 S41495	dc stretch-binding
23	197	5.6	464	2 S43363	transformation upr
24	194	5.5	362	2 S78515	single-stranded nu
25	193	5.5	464	2 A54143	kappa-B motif-bind
26	191	5.4	413	2 S46109	hnRNP complex prot
27	190.5	5.4	1220	2 T19117	hypothetical prote
28	190	5.4	1198	2 B88279	protein C08H9.2 li
29	182.5	5.2	1222	2 S56030	SCP160 protein - Y
30	178	5.0	381	2 S45766	hypothetical prote
31	165	4.7	806	2 T13690	hypothetical prote
32	164	4.6	784	2 J00317	hypothetical 82K p
33	162.5	4.6	848	2 S02262	glutenin high mole
34	161	4.6	891	2 G84693	probable proline-r
35	159	4.5	389	2 C96460	FI4M2.18 protein -
36	156.5	4.4	648	2 S04832	glutenin high mole
37	156.5	4.4	1283	2 T49692	related to SCP160
38	155.5	4.4	228	2 S42933	hypothetical prote
39	155.5	4.4	255	2 J00320	hypothetical 24.7K
40	153	4.3	1460	1 EDBE1F	immediate-early pr
41	153	4.3	1872	2 S36152	MHC class III hist
42	151.5	4.3	495	2 T04255	hypothetical prote
43	151.5	4.3	2715	2 T13049	eyelid - fruit fly
44	151	4.3	838	1 EEWTHW	glutenin, high mol
45	150	4.2	2142	2 B35098	MHC class III hist

ALIGNMENTS

RESULT 1

T23837 hypothetical protein M88.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T23837

R:Salston, J.

submitted to the EMBL Data Library, June 1994

A:Reference number: Z19806

A:Accession: T23837

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-854 <WIL>

A:Cross-references: EMBL:Z34802; PIDN:CAA84338.1; GSPDB:GN00021; CESP:M88.5

A:Experimental source: clone M88

C:Genetics:

A:Gene: CESP:M88.5

A:Map position: 3

A:Introns: 38/3; 96/1; 133/3; 255/3; 307/3; 395/2; 436/2; 492/3; 657/3; 759/1; 820/1

Alignment Scores:

Pred. No.:	5.99e-24	Length:	854
Score:	458.50	Matches:	160
Percent Similarity:	39.04%	Conservative:	91
Best Local Similarity:	24.88%	Mismatches:	212
Query Match:	12.98%	Indels:	180
DB:	2	Gaps:	21

US-09-270-437D-7 (1-1946) x T23837 (1-854)

QY	254	CAACCGGAGCAGACGAGCCATCATGAAGCTGAATGCCAC---CAGTTGGAGAA	310
DB	267	GlnAlaGlnHisGlnGlnSerHisGlnSerGln-AsnHisAsnGlnHisArgAs	286
QY	311	CCATGCCCTGAAGGTCCTCTACATCCCGCATGACAGATACACAGGAGCTCGAATGG	370
DB	286	nHis-----AsnGlnSerHisSerGlyProHisHis--	296


```
QY 873 GAGAAATTGTTGCGGCGCAGGAGGAGAAATATGAGAAAGTTTCGGAGGCGCTATGAGAAAT 932
Db ||| ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
334 GluThrCysAlaLysAlaGluGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
QY 933 GATGTGGCTGCCATGAGC-----TCTCACCTGATCCCTGGCCTGAACCTGGCTGCTGTA 986
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
354 AspIleAlaSerMetLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeu 373
QY 987 GGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCTCCAGCAGCGTT 1037
Db ||| ||| ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
374 GlyLeuPheProThrSerGlyMetProProThrSerGlyProProSerAlaMet 393
QY 1038 ACTGGGGCTGCTCCTTATGCTAGCTCCTTATGCTAGCTCCCGAGCAGGAGATGGTGCAGGTG 1097
Db ||| ||| ||| ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
394 Thr-----ProProTyrProGlnPheGlnGlnSer---GluThrGluThrValHisLeu 410
QY 1098 TTTATCCCGCGCCAGCGCGCGCCATCATCGCAAGAGGGGAGCAGACATCAAAACAG 1157
Db ||| ||| ||| ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyHisIleLysGln 430
QY 1158 CTCTCCCGGTTGCCAGCGCTCCATCAAGATTGCACACCCCGAAACACCTGACTCCAA 1217
Db ||| ||| ||| ||| ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY 1218 GTTCGTATGGTTATCATCTGACCGCCAGAGCGCCCAATTCAGGCTCAGGGAAGAATC 1277
Db ||| ||| ||| ||| ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY 1278 TATGGCAAACTCAAGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACC 1337
Db ||| ||| ||| ||| ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
471 TyrGlyLysIleLysGluLeuAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY 1338 CACATACGTGTGCCAGATCAGCAGCTGCGCGGTCAATTCGCAAAAGTGGAAAAACGGTG 1397
Db ||| ||| ||| ||| ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyGlyLysThrVal 510
QY 1398 AACGAGTTGCAGAAATTGACGGCAGCTGAGTGTAGTACCAAGACAGACAGCCCTGAT 1457
Db ||| ||| ||| ||| ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
QY 1458 GAGAACACCGAGTTCATCGTGAATAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAA 1517
Db ||| ||| ||| ||| ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
531 GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
QY 1518 CGGAAGATCCGAGACATCTCTGGCCCGAGGTTAACAG---CAGCATCAGAAAGGGA---CAG 1571
Db ||| ||| ||| ||| ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
QY 1572 AGTAACACGAGCCGAGCAGCGGAGAG 1598
Db ||| ||| ||| ||| ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||
571 SerGlyProProGlnSerArgLys 579
```

Search completed: July 23, 2004, 11:57:27

Job time : 119.853 secs

```
QY 753 AAGGAAGACCGAACCCTGAGAGAGGTAGAGCAAGATACCGAGACAAATAATCACCATTCTCC 812
Db 294 LysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
QY 813 TCGTGTGCAAGACCTTACCCTTTCAACCCCTGAGAGGACCATCACTGTGGAAGGGGCCATC 872
Db 314 ProLeuGlnGluLeuThrLeuTyAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY 873 GAGAAATTGTTCCAGGGCCGAGCAGGAAATAATGAGAAAGTTCCGGAGGCCCTATGAGAAAT 932
Db 334 GluThrCysAlaLysAlaGluGluGluIleMetLysLysIleArgGluSerTyrgluAsn 353
QY 933 GATGTGGCTGCATGAGC-----TCTCACTGATCCCTGGCTGGAACCTGGCTGCTGTA 986
Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeu 373
QY 987 GGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCNCCAGCAGCGCTT 1037
Db 374 GlyLeuPheProThrSerSerGlyMetProProThrSerGlyProProSerAlaMet 393
QY 1038 ACTGGGGCTGCTCCCTATAGCTCTTTATGAGGCTCCCGAGCAGAGATGGTGCAGGTG 1097
Db 394 Thr-----ProProTyProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
QY 1098 TTTATCCCGCCCGCAGGAGTGGGCCCATCATCGCAAGATTCGCAAGAGAGGGGCGACACATCAACAG 1157
Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGln 430
QY 1158 CTCTCCCGTTTGCAGCGCTCCATCAAGATTCGACCCAGCCGCAACACCTGACTCCCAA 1217
Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY 1218 GTTCGTATGGTTATCATCACTGGACCGCCAGAGGCCCAATTCAGGGCTCAGGAGAAATC 1277
Db 451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY 1278 TATGGCAAACTCAAGAGGAGAACTTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACC 1337
Db 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY 1338 CACATACGTGTGCCAGCATCAGCAGCTGCGGGTCAATTGGCAAGGTTGGAACGCTG 1397
Db 491 HisIleArgValProSerPheAlaGlyArgValIleGlyLysGlyGlyLysThrVal 510
QY 1398 AACGAGTTGCAGAAATTCAGCGCAGCTGAGTGTGTAGTACCAAGACAGCAGACCCCTGAT 1457
Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
QY 1458 GAGAACGACCGGTTCATCGTGAATAATCATCGACATTTCTATGCCAGTCCAGATGCTCAA 1517
Db 531 GluAsnAspGlnValValLysIleThrGlyHisPheTyAlaCysGlnValAlaGln 550
QY 1518 CGGAGATCCGAGACATCTGCGCCCGAGCTTAAGCAG---CAGCATCAGAGGGA---CAG 1571
Db 551 ArgLysIleGlnGluIleThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
QY 1572 AGTAACCGAGCGCCAGCGCAGGAGGAG 1598
Db 571 SerGlyProProGlnSerArgArgLys 579
```

RESULT 15

```
US-10-313-986-348
; Sequence 348, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, TongLong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
```

```
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313.986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-313-986-348
```

```
Alignment Scores:
Pred. No.: 1-52e-133 Length: 579
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best Local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 15 Gaps: 9
```

US-09-270-437D-7 (1-1946) x US-10-313-986-348 (1-579)

```
QY 102 CGGTTAGGAGCCGGAATAATTCAAATCGAAATATTCACCCAGCTCCGATGGGAAGTA 161
Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluVal 96
QY 162 CTGACAGCCCTGCTGGCTCAGTATGGTACAGTAGAAGACTGTGAGCAAGTGAACACCGAG 221
Db 97 LeuAspSerLeuLeuValGlnTyrglyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGAGAGCGGAGTGGTGAATGTCACTTATTCACCCGCGGAGCAGACCCAGGCAAGCCATC 281
Db 117 SerGluThrAlaValValAsnValThrTyrsSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATGAGCTGAATGGCCACCAGTGGAGAACCATGCTCCCTGAAAGTCTCTTACATCCCGAT 341
Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrlleProAsp 156
QY 342 GAGCAGATAGCA---CAGGGACCTGAGAAATGGGCCCGCGAGG-----GGCTTTGGCTCT 392
Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGln 176
QY 393 CGGGTTCAGCCCGCCGAGGCTCAGCTGTGGCAGCGGGCCCGCCAGCCAGCAAGCAGCAAA 452
Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY 453 GTGACATPCCCTTCGCTCCTGTGCTGCCACCCAGTATGTGGTGCCCATTTATGGCAAG 512
Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
QY 513 GAGGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCAGATAGACGTGCATAGG 572
Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY 573 AAGGAGAACGCGAGGTGAGCTGAAAGCCATCAGTGTGCTCCACCCCTCAGGCGCTGC 632
Db 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
QY 633 TCCTCCGCTGTGATGATGATCTTGAGATTTATGCATAAAGAGGCTTAAGACACCAAAACG 692
Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
QY 693 GCTGACGAGGTTCCCTGAAAGATCCTGGCCCATATAACTTTGTAGGGGCTCTCATGGC 752
Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY 753 AAGGAAGCAGCAACCTCAAGAGGTAGAGCAAGATACCGACACAAATAATCACCATCTCC 812
Db 294 LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
QY 813 TCGTTGCAAGACCTTACCCCTTTACACCCCTGAGAGGACCATCAGTGTGAAGGGGCCATC 872
Db 314 ProLeuGlnGluLeuThrLeuTyAsnProGluArgThrIleThrValLysGlyAsnVal 333
```

```
QY 693 GCTGACGAGTTCCCTGAGATCCCTGGCCCAATAAATCTTTGTAGGGCTCTCATTTGGC 752
Db 274 ThrGluGluLeuProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuLeuGly 293
QY 753 AAGGAGGCGGAACTGAGAGGTAGACAGATACCCAGACAAATAATCACATCTCC 812
Db 294 LysGluGlyArgAsnLeuLysIleGluGluAspThrAspThrLysIleThrIleSer 313
QY 813 TCCTTCAACACCTTACCTTTTCAACCTCTGAGAGGACCATCTACTGTGAAGGGGGCCATC 872
Db 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY 873 GAGAAATTGTCAGGGCCGAGCGGAGGAAATAATGAAGAAATTCGGAGGCGCTTAGAAT 932
Db 334 GluThrCysAlaLysAlaGluGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
QY 933 GATGTGGCTGCCATGAGC-----TCTCACCCTGATCCCTGGCTGACCTGGCTGTGTA 986
Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeu 373
QY 987 GGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCTCCAGCAGCGTT 1037
Db 374 GlyLeuPheProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
QY 1038 ACTGGGGCTCTCCTTAGCTCTTTATGAGCTCCCGAGGCTCCCGAGGAGATGGTGAGTG 1097
Db 394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
QY 1098 TTTATCCCGCCCGAGGCGGCGCCATCATCGCAAGAGCGGAGCAGCATCAACACAG 1157
Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlnHisIleLysGln 430
QY 1158 CTCCTCCGGTTGCCAGCGCTCCATCAGATTGCACACCCCGAAACACCATGCTCAAA 1217
Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY 1218 GTTCGTATGTTATCATCATGACCGCCGAGCGCCCAATTCAGGCTCAGGGAAGATC 1277
Db 451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY 1278 TATGGCAAACTCAGGAGGAGAACTCTTTGGTCCCAAGGAGGAGTGAAGCTGGAGACC 1337
Db 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY 1338 CACATAGCTGTCCAGCATCAGCAGCTGGCGGGTCAATTGGCAAAAGGTGAAAAACGGTG 1397
Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
QY 1398 AACGATTCAGAAATTTGCGGCGAGCTGAGTGGTAGTACCAAGACGACGACCTTGAT 1457
Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
QY 1458 GAGAAACGACGAGTCACTCGTGAATAATCATCGGACATTTCTATGCGCATCGATGGCTCAA 1517
Db 531 GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
QY 1518 CGGAAGATCCGAGACATCTCGGCCCGCCAGGTAAAGCAG---CAGCATCAGAGGGGA---CAG 1571
Db 551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
QY 1572 AGTAACCGCCCGCAGGACGAGGAG 1598
Db 571 SerGlyProProGlnSerArgLys 579
```

RESULT 14

```
US-10-117-982-480
; Sequence 480, Application US/10117982
; Publication No. US20030139438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
```

```
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 480
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-982-480
```

```
Alignment Scores:
Pred. No.: 1-52e-133 Length: 579
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best Local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 14 Gaps: 9
```

US-09-270-437D-7 (1-1946) x US-10-117-982-480 (1-579)

```
QY 102 CGGTTAGAGCGCGGAAAATTCAAATCCGAATATTCACCCCGAGCTCCGATGGGAAGTA 161
Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluVal 96
QY 162 CTCGACAGCTGCTGCTCAGTATGCTACAGTAGAAGTGTGAGCAAGTGTGAGCAAGTGTGAG 221
Db 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGACAGCGCAGTGTGTGATGTCCTATTCACCGCGGAGCAGACCGAGGCAAGGCATC 281
Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATGAAGCTGAATGCGCCACAGTGTGGAGAACCATGCCCTGAGGTCTCTACATCCCGCAT 341
Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY 342 GAGCAGATAGCA---CAGGACCTGAGAATGGGCGCGAGGG-----GGCTTTGGCTCT 392
Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGln 176
QY 393 CGGGGTGAGCCCGCAGGCGCTCCTGTGGAGCGCGGGGCGCCCGCAGCAGGAGCAGCAA 452
Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY 453 GTGGACATCCCTCTCGGCTCTGTGGTGGCCACCCAGTATCTGGTGGTCCATTATTGGCAAG 512
Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
QY 513 GAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCCAAGATAGACGTGCATAGG 572
Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY 573 AAGGAGACGCGAGTGCAGCTGAAAGGCCATCAGTGTGCTCCACCTGAGGGCTGC 632
Db 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
QY 633 TCCTCCGCTGTAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGACACCAAAACG 692
Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
QY 693 GCTGACGAGGTTCCTCGAAGATCTCGGCCCAATAAATCTTTAGGCGCTCTCATTTGGC 752
Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnPheValGlyArgLeuLeuGly 293
```

```

Db      254 SerAlaAlaCysLysSerIleLeuGluIleMethHisLysGluAlaGlnAsnIleLysPhe 273
Qy      693 GCTGACGAGTTCCTCCGAAAGTCTGCGCCATATAACTTTGTAGGCGCTCTATTGGC 752
Db      274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
Qy      753 AAGGAAGCAGCACTGGAAGAGTAGAGAATACCGAGACACAAAATCACCATCTCC 812
Db      294 LysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
Qy      813 TCGTTGCAAGACCTTACCTTTACACCTGAGGAGCACCATCTGTGAAGGGGCCATC 872
Db      314 ProLeuGlnGluLeuThrLeuTyraAsnProGluArgThrIleThrValGlyAsnVal 333
Qy      873 GAGAATTTCTGCGAGCGCAGCAGCAAAATATAAGAAAGTTTCGGGAGGCTTATGAGAA 932
Db      334 GluThrCysAlaLysAlaGluGluIleMethLysLysIleArgGluSerTyraGluAsn 353
Qy      933 GATGTGGTGCATGAGC-----TCTCACTGTATCTCTGGCTGAACTCTGGCTCTCTA 986
Db      354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeu 373
Qy      987 GGTCTTTTCCAGCTTCATCAGCGCAGCTCCGCG-----CCTCCGAGCAGCCTT 1037
Db      374 GlyLeuPheProProThrSerGlyMetProProThrSerGlyProProSerAlaMet 393
Qy      1038 ACTGGGGTGTCTCCCTATAGCTCTTTATGACGGCTCCGAGCAGGAGTGTGTGAGGTG 1097
Db      394 Thr-----ProProTyraProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
Qy      1098 TTTATCCCGCCGAGCAGTGGGCGCCATCATTCGCAAGAGAGGGCAGACATCAACAG 1157
Db      411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGln 430
Qy      1158 CTCTCCCGTTTGGCAGCGCTCCATCAAGATTGCAACCCGCAACACCTGACTCCAAA 1217
Db      431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
Qy      1218 GTTCGTATGTTTATCATCTGAGCGCAGCGCCAGAGGCCCAATTCAGGCTCAGGGAAGAATC 1277
Db      451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
Qy      1278 TATGCAAACTCAAGAGGAGAACTCTTTTGTCCCAAGAGGAAAGTGAAGCTGAGACC 1337
Db      471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
Qy      1338 CACATACGTGTGCCAGCATCAGCAGCTGCGCGGTTCATTGGCAAGGTGGAAACGCTG 1397
Db      491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
Qy      1398 AACGAGTTGCAGAAATTTACGCGCAGCTGAGGTGGTAGTACCAAGAGACACACCCCTGAT 1457
Db      511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
Qy      1458 GAGAACGACAGGTCATCGTGAAATCATTCGACATTTCTATGCCAGTCAGATGCTCAA 1517
Db      531 GluAsnAspGlnValValLysIleThrGlyHisPheTyraLysGlnValAlaGln 550
Qy      1518 CGGAAGATCCGAGACATCTCGCCAGGTAAAGCAG---CAGCATCAGAAGGGA---CAG 1571
Db      551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
Qy      1572 AGTAACACGAGCGCCAGGACGAGGAGAG 1598
Db      571 SerGlyProProGlnSerArgArgLys 579

```

RESULT 13

US-10-117-982-449
; Sequence 449, Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.

```

; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriack
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; OF LUNG CANCER
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; SOFTWARE: FASTSEQ for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 484
; SEQ ID NO 449
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-982-449

Alignment Scores:      1.52e-133      Length:      579
Pred. No.:      1875.00      Matches:      380
Score:      84.68%      Conservative:      51
Best Local Similarity:      74.66%      Mismatches:      62
Query Match:      53.10%      Indels:      16
DB:      14      Gaps:      9
US-09-270-437D-7 (1-1946) x US-10-117-982-449 (1-579)
Qy      102 CGGTTAGGAGCGGAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTA 161
Db      77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluVal 96
Qy      162 CTGACAGCCTGTGGCTCAGTATGGTACAGTAGAAGTGTGACCACTGTGACCACTGACACCCGAG 221
Db      97 LeuAspSerLeuLeuValGlnTyrglyValGluSerCysGluGlnValAsnThrAsp 116
Qy      222 AGTGAGCGGAGTGGTGAATGTCACTATTCCACCGGAGCAGACCCAGGCAAGCCATC 281
Db      117 SerGluThrAlaValAlaValAsnValThrTyrSerLysAspGlnAlaArgGlnAlaLeu 136
Qy      282 ATGAGCTGAATGGCCACAGTTGGAGAACCATGCTCCCTCAAGGTCTCTACATCCCGAT 341
Db      137 AspLysLeuAsnGlyPheGluLeuGluAsnPheThrLeuLysValAlaIleIleGlyLys 156
Qy      342 GAGCAGATAGCA---CAGGAGCTGAGAAATGGGCGCCGAGG-----GGCTTGGCTCT 392
Db      157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGln 176
Qy      393 CGGGTTCAGCCCGCCGAGGCTCACCTGTGGCAGCGGGCCCGCCAGCAGCAGCAGCAA 452
Db      177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
Qy      453 GTGACATCCCTTCGGCTCTCTGTGTCGCCACCCAGTATGTGGTGCCATTTATGGCAAG 512
Db      194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
Qy      513 GAGGGGCGCACCATCCGCAACATCAAAACAGACCCAGTCCAGATAGACCTGATAGG 572
Db      214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
Qy      573 AAGGAGAACGAGGTGAGCTGAAAGCCATCAGTGTGCATCCACCCCTGAGGCTGC 632
Db      234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
Qy      633 TCCTCCGCTTGAAGATGATCTTGAGATTATGCATAAAGAGGTGAAGACACCAAG 692
Db      254 SerAlaAlaCysLysSerIleLeuGluIleMethHisLysGluAlaGlnAspIleLysPhe 273

```

Db 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
 QY 633 TCCTCCGCTGTAGATGATCTTGGAGATTATGATTAAGAGCGCTAAGGACACCAAAACG 692
 Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
 QY 693 GCTGACGAGGTCCCTCGAGATCCTGCGCCATATTAACCTTGTAGGCGCTCATGGC 752
 Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
 QY 753 AAGGAAGCGAGAACCTGAAGAGGTAGACAGATACCGAGACAAATAACCATCTCC 812
 Db 294 LysGluGlyArgAsnLeuLysIleLeuGlnAspThrAspThrLysIleThrIleSer 313
 QY 813 TCCTTGAAGACCTTACCTTTACACCTGAGAGGACCATCACTGTGAAGGGGGCCATC 872
 Db 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
 QY 873 GAGAATTGTTGCGAGCGGACGAGGAATAATGAAGAAAGTTCGGAGGCGCTATGAGAAT 932
 Db 334 GluThrCysAlaLysAlaGluGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
 QY 933 GATGTGCTGCCATGAGC-----TCTCACCTGATCCTCGCTGCAACCTGGCTGTGTA 986
 Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisIleLeuProGlyLeuAsnLeuAlaLeu 373
 QY 987 GGTCTTTCCAGCTTCATCCAGCGCAGTCCGCG-----CCTCCAGCAGCGTT 1037
 Db 374 GlyLeuPheProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
 QY 1038 ACTGGGGCTGCTCCTATAGCTCTTATCAGGCTCCCGAGCAGGAGATGTCAGGTG 1097
 Db 394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
 QY 1098 TTATATCCCGCCAGGCGAGTGGCGCCATCATCGCAAGAGAGGCGAGCACATCAACAG 1157
 Db 411 PheIleProAlaLeuSerValGlyAlaIleLeuGlyLysGlnGlyHisIleLysGln 430
 QY 1158 CTCTCCCGGTTGGCAGCGCTCCATCAAGATTGCACCCGAGAACACCTGATCTCAAA 1217
 Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
 QY 1218 GTTCGTATGTTATCATCTGACCGCCAGCGCCCAATTCAAGGCTCAGGAGGAATC 1277
 Db 451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
 QY 1278 TATGGCAAACTCAGGAGGAGAACTTCTTTGGTCCCAAGAGGAGTGAAGCTGGAGAC 1337
 Db 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
 QY 1338 CACATACGTGTCAGCATCAGCAGCTGCGGGGTGATTGGCAAGGTGGAATAACGGTG 1397
 Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
 QY 1398 AACGAGTTGAGATTTGACGCGAGCTGAGTGTAGTACCAAGAGACGACCCCTGAT 1457
 Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
 QY 1458 GAGAAGACCGAGTCACTGCGAAATCATCCGACATTTCTATCCAGTCAAGTGGCTCAA 1517
 Db 531 GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
 QY 1518 CGGAAGATCCGAGACATCTCGCCCGGTTAAGCAG---CAGCATCAGAGGGA---CAG 1571
 Db 551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnValAlaLeuGln 570
 QY 1572 AGTAACCGAGCGCCAGCAGGAGAG 1598
 Db 571 SerGlyProGlnSerArgArgLys 579

RESULT 12
 US-10-117-982-446

; Sequence 446, Application US/10117982

; Publication No. US20030138438A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FOY, Teresa M.
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Vedvick, Thomas S.
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Watanabe, Yoshihiro
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Mericle, Barbara
 ; APPLICANT: Spies, Gregory A.
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Wang, Tongtong
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.455C18
 ; CURRENT APPLICATION NUMBER: US/10/117,982
 ; CURRENT FILING DATE: 2002-04-05
 ; NUMBER OF SEQ ID NOS: 484
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 446
 ; LENGTH: 579
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-117-982-446

Alignment Scores:
 Pred. No.: 1,52e-133 Length: 579
 Score: 1875.00 Matches: 380
 Percent Similarity: 84.68% Conservative: 51
 Best Local Similarity: 74.66% Mismatches: 62
 Query Match: 53.10% Indels: 16
 DB: 14 Gaps: 9

US-09-270-437d-7 (1-1946) x US-10-117-982-446 (1-579)

QY 102 CGGGTAGAGCGCGGAAATTCGAATCCGNAATATTCACCCCGAGCTCCGATGGGAAGTA 161
 Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProHisLeuGlnTyrGluVal 96
 QY 162 CTGACAGCCTGCTGCTCAGTATGTACAGTAGAGAATGTGTCAGCAAGTGAACACCGAG 221
 Db 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
 QY 222 ATGAGACGGCAGTGTGTGATGTCACTATTCCACCGGAGCAGACCGAGCGAAGCATC 281
 Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
 QY 282 ATGAAGCTGAATGGCCACCGAGTTGGAGAACCATGCCCTGAGGCTCTCTACATCCCGAT 341
 Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
 QY 342 GAGCAGATAGCA---CAGGGACCTGAGAATGGCGCCGAGGG---GGCTTTGGCTCT 392
 Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGln 176
 QY 393 CGGGGTGACCCCGCAGGGCTCACCTGTGGCAGCGGGGGCCCGACCGAAGCAGCAGCAA 452
 Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
 QY 453 GTGGACATCCCTCTCGGCTCCTGGTGGCCACCGCAGTATGTGGTGGCTCCATTTGGCAAG 512
 Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
 QY 513 GAGGGGGCCACCATCCCGCAACATCAAAACAGACCCAGTCCCAAGATAGACGTGCATAGG 572
 Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
 QY 573 AAGGAAACGCGAGGTGCAGCTGAAAAGCCATCAGTGTGCATCTCCACCCCTGAGGCTGC 632
 Db 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
 QY 633 TCCTCCGCTTGTAAAGTATGATCTTGGAGATTATGCATAAAGAGGCTTAAGACACCAAAACG 692

RESULT 11
 US-10-117-982-348
 ; Sequence 348, Application US/10117982
 ; Publication No. US2003013438A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FOV, Teresa M.
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Vedvick, Thomas S.
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Watanabe, Yoshihiro
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Maricle, Barbara
 ; APPLICANT: Spies, Gregory A.
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Wang, Tongtong
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.455C18
 ; CURRENT APPLICATION NUMBER: US/10/117,982
 ; CURRENT FILING DATE: 2002-04-05
 ; NUMBER OF SEQ ID NOS: 484
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 348
 ; LENGTH: 579
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-117-982-348

Alignment Scores:
 Pred. No.: 1,52e-133 Length: 579
 Score: 1875.00 Matches: 380
 Percent Similarity: 84.68% Conservative: 51
 Best Local Similarity: 74.66% Mismatches: 62
 Query Match: 53.10% Gaps: 16
 DB: 14

US-09-270-437D-7 (1-1946) x US-10-117-982-348 (1-579)
 QY 102 CGGTTAGGAGCGGAAATTCAGAAATATCCACCCAGCTCCGATGGAAAGTA 161
 Db 77 ArgGlnArgIleArgIysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluVal 96
 QY 162 CTGACAGCGCTGCTGCCTCAGTATGTGTACAGTATAGAGAACTGTGAGCAAGTGAACACCGAG 221
 Db 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
 QY 222 AGTCAGAGCGGAGTGTGTAATGTACCTATTCACACCGGAGGAGCAGCCAGCGCATC 281
 Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
 QY 282 ATGAAGCTGAATGGCCACCACTGTTGGAGAACATCCCTGAGAGTCTCTATCATCCCGAT 341
 Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
 QY 342 GAGCAGATAGCA---CAGGGACCTGAGAAATGGGGCGCGAGGG-----GGCTTGGCTCT 392
 Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGln 176
 QY 393 CGGGGTGAGCCCGCCAGGCTACCTGTGGCAGCGGGGGCCCGACCGAGCAGCAGCAA 452
 Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
 QY 453 GTGACATATCCCTTCCTGGCTCTGGTGCACCCAGTATGTGGGTGCATATTATGGCAAG 512
 Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleGlyLys 213
 QY 513 GAGGGGCGCCATCCCAATCATCAAAACAGACCCAGTCCCAAGATAGAGTGCATAGG 572
 Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
 QY 573 AAGGAGAACGCGAGTGCAGCTGAAAGAACCATCATAGTGTGCACCTCCACCTTGGGGTGC 632

```
Db 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
QY 633 TCCTCCGCTTGTAAAGATGATCTTGGAGATTATGATAAAGAGGCTAAAGACACCAAAACG 692
Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
QY 693 GCTCAGCAGGTTCCCTCAGATCCTGCCCATATAACTTTGTAGGCGCTCATTTGGC 752
Db 274 ThrGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY 753 AAGGAAGCAGCAACCTGAAGGTAGAGCAAGATACCGAGACAAAAATCACCATCTCC 812
Db 294 LysGluGlyArgAsnLeuLysLysIleGluGlnAspThrLysIleThrIleSer 313
QY 813 TCGTGTCAAGACCTTACCTTTACAACCTGAGAGGACCATCACTGTGAAGGGGCCCATC 872
Db 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY 873 GAGATTTGTCAGGCGGAGCAGCAAGAAATATGAAGAGTTCGGGAGGCTATGAGAT 932
Db 334 GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
QY 933 GATGTGCTGCATGAGC-----TCTCACCTGATCCTGCTGCTGAACCTGGCTGCTGTA 986
Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeu 373
QY 987 GGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CTCCAGCAGCGTT 1037
Db 374 GlyLeuPheProThrSerGlyMetProProThrSerGlyProProSerAlaMet 393
QY 1038 ACTGGGCTGCTCCTATAGTCTCTTATCAGCTCCCGAGGAGGATGTCAGGTC 1097
Db 394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
QY 1098 TTTATCCCGCCAGCAGGTGGCGCCATCATCCGCAAGAGGGCGAGCATCAAAACAG 1157
Db 411 PheIleProAlaLeuSerValGlyAlaIleGlyLysGlnGlyGlnHisIleLysGln 430
QY 1158 CTCTCCGTTTGCAGCGCTCCATCAAGATTGCACCCGCAACCTGATCCCAA 1217
Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY 1218 GTTGTGTGTATCATCACTGACCGCCAGAGCCCAATTCAGGCTCAGGAGGATC 1277
Db 451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY 1278 TATGGCAACTCAAGAGGAGAACTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGCC 1337
Db 471 TyrGlyLysIleLysGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY 1338 CACATAGTGTCCAGCATCAGAGCTGGCGGGTCAATGGCAAGGTGGAANAACGGTG 1397
Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
QY 1398 AACGAGTTGCAGATTTCAGCGAGCTGAGTGTAGTACCAAGACAGACCAACCCCTGAT 1457
Db 511 AsnGluLeuGlnAsnLeuSerAlaGluValValProArgAspGlnThrProAsp 530
QY 1458 GAGAACACCAAGGTTCATCTGTGAATCATCGACATTTCTATGCACTCAGATGGCTCAA 1517
Db 531 GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
QY 1518 CGGAGATCCGAGACATCTGGCCCGAGTTAAGCAG---CAGCATCAGAGGGA---CAG 1571
Db 551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
QY 1572 AGTAACCAAGGCGCCAGGACGAGGAAG 1598
Db 571 SerGlyProProGlnSerArgLys 579
```

RESULT 10

US-10-007-700-449

; Sequence 449, Application US/10007700

```
; Publication No. US20030064947A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Li, Samuel X.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckman, David W.
; APPLICANT: Cai, Peng
; APPLICANT: Poi, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C17
; CURRENT APPLICATION NUMBER: US/10/007,700
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 449
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-007-700-449

Alignment Scores:
Pred. No.: 1,52e-133 Length: 579
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best Local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 12 Gaps: 9

US-09-270-437D-7 (1-1946) x US-10-007-700-449 (1-579)
```

```
QY 102 CGGGTAGAGAGCGGAAATTCAAATCCGAAATATTCACCCAGCTCCATGGGAAGTA 161
Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluVal 96
QY 162 CTGGACAGCCTGCTGGCTCAGTATGTACATAGAGAACTGTGAGCAAGTGAACACCCGAG 221
Db 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGAGACGGCAGTGTGAATGCACCTATTCCAAACGGGAGAGACCCAGGCAAGCATC 281
Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATGAGCTGAATGCCACACAGTTGGAGAACCATGCCCTGAGGCTCTCTACATCCCGAT 341
Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY 342 GAGCAGATAGCA---CAGGGACCTGAGAAATGGGCGCGAGGG-----GGCTTTGGCTCT 392
Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgargGlyLeuGlyGln 176
QY 393 CGGGGTGAGCCCGCAGGCTCACCTGTGGGAGCGGGGCCCGCCAGAGAGAGAGCAA 452
Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY 453 GTGAGATCCCTCTGGCTCTGCTGGTGGCCACCGATGTGGTGGCTGCTATTTGGCAAG 512
Db 194 CysAspLeuProLeuArgLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
QY 513 GAGGGGCGCCACCATCCGCAACATCAAAAACAGACCCAGTCCCAAGATAGAGTGCATAGG 572
Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
```

```

||||:|||||
254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
QY GCTGACGAGGTTCCCTCAAGATCCTGCGCCATATAACTTTGTAGGGCGTCTCATTCGGC 752
Db ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY AAGGAAGACCGAAGCTCAAGAGGTAGACGAGATACCGACAAATAATCACATCTCC 812
Db LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
QY TCGTTGCAAGACCTTACCTTTACACCTGAGAGGACCATCTGTAAGGGGCCATC 872
Db ProLeuGlnGluLeuThrLeuTyraAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY GAGAATTTGTCAGGCGCCGACGAGGAATAATGAAGAAATTCGGGAGCGCTTATCAGAA 932
Db GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyroGluAsn 353
QY GATGTCGTCGATGAGC-----TCTCACTGTATCCTGCTGAGCTGAACTGCTGCTGTA 986
Db AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeu 373
QY GGTCTTTTCCAGCTTCATCAGCCAGCTCCGCG-----CCTCCGAGCAGCGTT 1037
Db GlyLeuPheProProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
QY ACTGGGGTGCTCCCTATAGCTCCTTTATGAGGCTCCGAGCAGGATGTCAGGTG 1097
Db Thr-----ProProTyroProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
QY TTTATCCCGCCGAGGAGTGGCGCCATCATTCGCAAGAGAGGGCAGACATCAAAACAG 1157
Db PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyHisIleLysGln 430
QY CTCTCCCGGTTGCGAGCGCTTCATCAGATTGACACCCGCAACACCTGACCTCCAAA 1217
Db LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY GTTCTGTATGTTTATCATCTGACGCGCCGAGGCGCCCAATTCAGAGCTCAGGGAAGAATC 1277
Db ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY TATGCCAAACTCAAGGAGAGAACTTCTTTGTTCCCAAGGAGAAAGTGAAGCTGAGACC 1337
Db TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY CACATACGTGTCGACATCAGCAGCTGCGCGGTCATTGCGCAAGGTGGAAGGCGGTG 1397
Db HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
QY AACGAGTTGCAGAAATTCAGCGCAGCTGAGGTGGTAGTACCAAGAGACCCAGCCCTGAT 1457
Db AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
QY GAGAACGACGAGTTCATCGTGAATATCATCGACATTTCTATCCAGTCCAGATGCTCAA 1517
Db GluAsnAspGlnValValLysIleThrGlyHisPheTyraLysAlaCysGlnValAlaGln 550
QY CGAAGATCCGAGACATCTGCGCCAGGTAGCAG-----CAGCATCAGAGGGA---CAG 1571
Db ArgLysIleGlnGluIleuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
QY AGTAAACGAGCGCCGAGCGAGGAG 1598
Db SerGlyProProGlnSerArgLys 579

```

RESULT 9

US-10-007-700-446
; Sequence 446, Application US/10007700
; Publication No. US20030064947A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong

```

; APPLICANT: Wang, Aijun
; APPLICANT: Skelky, Yasir A.W.
; APPLICANT: Li, Samuel X.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckman, David W.
; APPLICANT: Cai, Feng
; APPLICANT: Foy, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C17
; CURRENT APPLICATION NUMBER: US/10/007,700
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 446
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-700-446

Alignment Scores:
Pred. No.: 1,52e-133 Length: 579
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best Local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 12 Gaps: 9

US-09-270-437D-7 (1-1946) x US-10-007-700-446 (1-579)
QY 102 CGGGTTAGGACCGCGAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTA 161
Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluVal 96
QY 162 CTGGACAGCTCTGGCTCAGTATGTCAGTACAGTACGAGTGTGAGCACTGTGACCACTGACACCGAG 221
Db 97 LeuAspSerLeuLeuValGlnTyroGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGAGACGCGAGTGTGAATGTCACTTATTCACCGGAGCAGACACGAGCAAGCCATC 281
Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATGAAGCTGAATGGCCACCGATTGGAGAACCATGCGCTGAGAGTCTCTACATCCCGAT 341
Db 137 AspLysLeuAsnGlyPheGlnLeuGlnAsnPheThrLeuLysValAlaTyroIleProAsp 156
QY 342 GAGCAGATAGCA---CAGGACCTGAGATGGCGCCGAGGG-----GGCTTGGCTCT 392
Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGln 176
QY 393 CGGGGTGACGCGCCGCGAGGGCTCACCTGTGGCAGCGGGGCGCCGAGCAGACAGCAAA 452
Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY 453 GTGGACATCCCTTGGCTCTGGTGTGCCACCCAGTATGTGGGTGCCATTTATTTGGCAAG 512
Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
QY 513 GAGGGGGCCACCATCGGAACATCACAAACAGACCCAGTCCCAAGATGACGTGCATAGG 572
Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY 573 AAGGAGAACGCGAGGTGACGCTGAAAAGCCATCAGTGTGCTACCTCCACCCCTGAGGCTGC 632

```



```
QY 693 GCTGACAGAGTTCCCTGAGATCTCTGGCCCAATAAATCTTTGTAGGGCTCTCATTTGGC 752
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY 753 AAGGAGGACGGACCTGAGAGAGGTAGACAGATACCGAGACAAAAATCACCATCTCC 812
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
294 LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
QY 813 TCCTTGCAGACCTTACCCCTTTACACCTCGAGAGGACCATCACTGTGAAGGGGGCCATC 872
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY 873 GAGAAATGTTGCAGGCGGACGAGGAATAATGAGAAAGTTCCGAGGCGCTATGAGAAAT 932
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
334 GluThrCysAlaLysAlaGluGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
QY 933 GATGTGGCTGCCATGAGC-----TCTCACCTGATCCCTGGCTGAGACCTTGGCTGTGTA 986
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
354 AspIleAsnSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeu 373
QY 987 GGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCTCCAGCAGCGTT 1037
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
374 GlyLeuPheProThrSerGlyMetProProThrSerGlyProProSerAlaMet 393
QY 1038 ACTGGGGCTCTCCCTATAGCTCTTTATCAGGCTCCCGAGCGAGATGGTGCAGGTG 1097
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
QY 1098 TTTATCCCGCCCGAGGCGGCGCCATCATCGGCAAGAGGGGAGCACATCAAAACAG 1157
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
411 PheIleProAlaLeuSerValGlyAlaIleIleGlyGlnGlyGlnHisIleLysGln 430
QY 1158 CTCTCCCGGTTTCCAGCGCTCCATCAAGATTGCACACCCGAAACACCTGACTCCAAA 1217
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY 1218 GTTCGTATGTTATCATCTGACCGCCAGGCGCCCAATTCAGGCTCAGGGAGAAATC 1277
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY 1278 TATGGCAAACTCAAGGAGAGAACTCTTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACC 1337
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY 1338 CACATAGCTGTCAGCATCAGCAGCTGGCGGCTCATTTGGCAAAAGTGAATAACGGTG 1397
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
QY 1398 AACGAGTTGCAGAAATTTGACGCGAGCTGAGTGTAGTACCAAGACACGACCCCTGAT 1457
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
QY 1458 GAGAAGCAGCAGGTTCATCGTGAATAATCATCGGCATTTCTATGCGAGTCAGATGGCTCAA 1517
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
531 GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
QY 1518 CGGAAGATCCGAGACATCTCGGCCCGAGTTAAGCAG---CAGCATCAGAGGGA---CAG 1571
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
551 ArgLysIleGlnIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
QY 1572 AGTAACCGAGCCCGAGGCGAGGGAAG 1598
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
571 SerGlyProProGlnSerArgLys 579
```

RESULT 8

```
US-10-007-700-348
; Sequence 348, Application US/10007700
; Publication No. US20030064947A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Li, Samuel X.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Beckman, David W.
; APPLICANT: Cai, Peng
; APPLICANT: Foy, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C17
; CURRENT APPLICATION NUMBER: US/10/007,700
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-700-348

Alignment Scores:
Pred. No.: 1,52e-133 Length: 579
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best Local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 12 Gaps: 9

US-09-270-437D-7 (1-1946) x US-10-007-700-348 (1-579)

QY 102 CGGGTAGAGCGCGGAAATTCAAATCCGAAATATTCACCCCGAGCTCCGATGGGAGTA 161
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluVal 96
QY 162 CTGGACAGCCTGCTGCTCAGTATGTGACAGTGTGAGAACTGTGAGCAAGTGAACACCGAG 221
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGlnGlnValAsnThrAsp 116
QY 222 AGTGACAGCGCAGTGTGATGTACCTATTTCACCGGAGCAGACCGAGCAAGCCATC 281
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATGAAGCTGAATGCCACCACTGTGGAGAACCATGCCCTGAAGGCTCTCCTACATCCCCGAT 341
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY 342 GAGCAGATAGCA---CAGGGACCTGAGAAATGGGCGCGCAGGG-----GGCTTTGGCTCT 392
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGln 176
QY 393 CGGGGTGACCGCCCGCAGGCGCTCACCTGTGGAGCGGGGCCCGCCAGCAGAGCGACAA 452
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY 453 GTGGACATCCCCCTCGGCTCCTGTGTCGCCACCCAGTATGTGGGTGCCATTATTGGCAAG 512
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
QY 513 GAGGGGGCCACCATCCGCAACATCAAAAACAGACCCAGTCCCAAGATAGACGTGCATAGG 572
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY 573 AAGGAGACCGAGGTGCAGCTGAAAAGCCATCAGTGTGCATCTCCACCCCTGAGGGCTGC 632
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
QY 633 TCCTCCGCTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGACACCAACG 692
```

633 TCCTCCGCTTGAAGATGACTTGGAGATTATGATTAAGAGGCTTAAGACACCAAAACG 692
 254 SerAlaAlaCysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
 693 GCTGACAGGTTCCCTGAAATCCTGCGCCATTAATTAATCTTTAGGGCGTCTCATGGC 752
 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
 753 AAGGAGACGACCACTCAAGAGGTAGACAGATACGAGACACAAATACCACTATCC 812
 294 LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
 813 TCGTTGCAAGACTTACCTTTACACCTGAGAGCACATCATCTGTAAGGGGCCATC 872
 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
 873 GAGAAATGTTGCGAGCGGACGAGCAAGAAATATGAAAGTTCGGAGGCGCTATGAGAT 932
 334 GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
 933 GATGTGGCTGCATGAGC-----TCTCACTGATCCCTGGCTCAACCTGGCTCTGTA 986
 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeu 373
 987 GGTCTTTTCCAGCTTTCATCCAGCGCAGTCCCGCG-----CCTCCAGCAGCGTT 1037
 374 GlyLeuPheProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
 1038 ACTGGGGCTGCTCCCTTAGCTCTTTATGAGGCTCCGAGCAGGATGGTGGAGGTG 1097
 394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
 1098 TTTATCCCGCCGCGAGTGGCGCCATCATCGCAAGAGAGGGCGAGCACATCAACAG 1157
 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyHisIleLysGln 430
 1158 CTCTCCCGGTTTTCAGCGCTTCATCAAGATTGACCAACCGCCGAAACCTCATCCCAA 1217
 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
 1218 GTTCTATGTTATCATCATCTGAGCGCGCAGAGGCCCAATTCAGGCTCAGGAGAAATC 1277
 451 ValArgMetValIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 470
 1278 TATGCAAACTCAAGAGAGAACTTCTTTGCTCCCAAGGAGAAAGTGAAGCTGGAGAC 1337
 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
 1338 CACATACGTGTGCCAGCATCAGCAGCTGCGCGGTTCATTGGCAAGAGTGGAACCGGTG 1397
 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
 1398 AACGATTCAGAAATTGCGCAGCTGAGGTGTAGTACCAAGAGACAGACCCCTGAT 1457
 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
 1458 GAGAACGACGACGCTCATCTGAAAATCATCGACATTTCTATGCCAGTCCAGATGGCTCAA 1517
 531 GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
 1518 CGAAGATCCGAGACATCTGCGCCAGGTAAAGCAG---CAGATCAGAGGGA---CAG 1571
 551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
 1572 AGTAACCGCCGACGAGGAG 1598
 571 SerGlyProGlnSerArgArgLys 579

RESULT 7

US-09-897-778-449

; Sequence 449, Application US/09897778

; Patent No. US20020147143A1

GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Marnerakis, Margarita
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Vedwick, Thomas S.
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Watanabe, Yoshihiro
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Peckham, David W.
 ; APPLICANT: Fanger, Neil
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.455C16
 ; CURRENT APPLICATION NUMBER: US/09/897,778
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 449
 ; LENGTH: 579
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-897-778-449

Alignment Scores:
 Pred. No.: 1,52e-133 Length: 579
 Score: 1875.00 Matches: 380
 Percent Similarity: 84.68% Conservative: 51
 Best Local Similarity: 74.68% Mismatches: 62
 Query Match: 53.10% Indels: 16
 DB: 9 Gaps: 9

US-09-270-437D-7 (1-1946) x US-09-897-778-449 (1-579)

QY 102 CGGTTAGAGCGCGAAATTCAAATCCAAATATTCACCCCGAGCTCGATGGAGATGA 161
 Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluVal 96
 QY 162 CTGGACAGCTGCTGGCTCAGTATGTACAGTAGAGACTGTGACAGTGTGACACCGAG 221
 Db 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
 QY 222 AGTGAGAGCGGAGTGGTGAATGTCACTATTCCAAACCGGAGCAGACAGGCAAGCCATC 281
 Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
 QY 282 ATGAAGCTCAATGGCCACCGAGTGGAGAACCAATGCCCTGAAAGTCTCTTACATCCCGAT 341
 Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
 QY 342 GAGCAGATAGCA---CAGGGACCTGAGAAATGGCGCGCGAGG-----GGCTTGGCTCT 392
 Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGln 176
 QY 393 CGGGGTGAGCCCGCGCGGCTCCTGTGGCAGCGGGGGCGCCAGCCAGCAGCAGCAA 452
 Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
 QY 453 GTGGACATCCCGCTTGGCTCTGTGGTCCCAACCCAGTATGTGGTGGCTTATTTGGCAAG 512
 Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
 QY 513 GAGGGGGCCACCATCCGACATCACAACAGACCCAGTCCCAAGATAGCATGCGCATAGG 572
 Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
 QY 573 AAGGAGACGCGAGGTGCGAGCTGAAAAAGCATCACTGTGCACTCCACCCCTGAGGCTGC 632
 Db 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
 QY 633 TCCTCCGCTTGAAGATGATCTTGAGATTATGATTAAGAGCTTAAGACACCAAAACG 692
 Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273


```
Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
QY 513 GAGGGGCGCCATCCGCAACATCACAAACAGACCCAGTCCACAGATAGAGTGCATAGG 572
Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY 573 AAGGAGAACGAGTGCAGTGTGAAGAACCCATCAGTGTGCTCCACCCCTGAGGCGTGC 632
Db 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrLysLeuSerThrProGluGlyThr 253
QY 633 TCCTCCGCTTGTAGATCATCTTGGAGTATTGATTAAGAGGCTTAAGGACACCAACAG 692
Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
QY 693 GCTCAGCAGGTTCCCTGAAATCTCTGGCCCATATAACTTTGTAGGGCGTCTCATTTGGC 752
Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY 753 AAGGAGGACGAACTGAAAGAGTGAAGCAAGATACGAGATACGAGACAAAATCACCATCTCC 812
Db 294 LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
QY 813 TCGTTGCAAGACTTACCTTTACACCTGAGGACCATCAGTGTGAGGGGCGCCATC 872
Db 314 ProLeuGlnGluLeuThrLeuTyAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY 873 GAGAATTTGTCAGGCGGACGAGAAATATGAAGAAAGTTGGGAGGCGCTTATGAGAA 932
Db 334 GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyArgLys 353
QY 933 GATGTGGTGCATGAGC-----TCTCAGCTGATCCCTGGCTGAACCTGGCTCTCTTA 986
Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeu 373
QY 987 GGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCTCCAGCAGCGTT 1037
Db 374 GlyLeuPheProProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
QY 1038 ACTGGGGTGTCTCCTATAGCTCTTATGAGGTCCTCCGAGGCGAGGATGTGTCAGGTG 1097
Db 394 Thr-----ProProTyProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
QY 1098 TTTATCCCGCCGAGCAGTGGCGCCATCATCGCAAGAGAGGCGAGCACATCAACAG 1157
Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyHisIleLysGln 430
QY 1158 CTCTCCCGTTTGCAGCGCTCCATCAAGATTGCAACCCGCAACACCTGACTCCAAA 1217
Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY 1218 GTTCGTATGTTTATCATCTGACCGCGCAGAGGCCCAATTCAGGCTCAGGGAAGAATC 1277
Db 451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY 1278 TATGCCAACTCAAGAGAGAACTCTTTGGTCCCAAGAGAGAGTGAAGCTGGAGACC 1337
Db 471 TyrGlyLysIleLysGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY 1338 CACATACGTGTCCAGCATCAGCAGCTGCGCGGTTCATTGGCAAGAGTGGAAAACGGTG 1397
Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
QY 1398 AACGAGTTGCAGAAATTTGACGCGAGCTGAGGTGGTAGTACCAAGAGACACAGCCCTGAT 1457
Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
QY 1458 GAGAACGACAGGTGCATCGTGAATAATCATCGACATTTCTATGCCAGTCAAGTGCCTAA 1517
Db 531 GluAsnAspGlnValValValLysIleThrGlyHisPheTyAlaCysGlnValAlaGln 550
QY 1518 CGRAGATCCGAGACATCTGCGCCAGGTTAAGCAG---CAGCATCAGAGGGA---CAG 1571
```

```
Db 551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
QY 1572 AGTAACAGAGCCCGACGACGAGGAGAG 1598
Db 571 SerGlyProProGlnSerArgArgLys 579
RESULT 5
US-09-897-778-348
; Sequence 348, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Warnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-778-348
Alignment Scores:
Pred. No.: 1,52e-133 Length: 579
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best Local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 9 Gaps: 9
US-09-270-437D-7 (1-1946) x US-09-897-778-348 (1-579)
QY 102 CGGTTAGAGCGCGGAAATTCAAATCCGAATATTCACCCAGCTCCGATGGGAAGTA 161
Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluVal 96
QY 162 CTGGACAGCTGCTGGCTCAGTATGCTAGTAGAGAACTGTGACAGTGTGACACCGAG 221
Db 97 LeuAspSerLeuLeuValGlnTyGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGACGCGCAGTGGTGAATGTCTACTATTCCAAACCGGGAGCAGACGAGGAGCCATC 281
Db 117 SerGluThrAlaValValAsnValThrTySerSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATGAGCTGAATGGCCACAGTTGGAGAACCATGCTCCCTGAGTCTCTACATCCCGAT 341
Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrrIleProAsp 156
QY 342 GAGCAGATAGCA---CAGGGACCTGAGAAATGGCGCCGAGGG-----GGCTTTGGCTCT 392
Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGln 176
QY 393 CGGGTTCAGCCCGCCAGGCTCAGCTGTGGCAGCGGGGGGCCCGCCAGCAGCAGCAA 452
Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY 453 GTGGACATCCCTTCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 512
Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
QY 513 GAGGGGCGCCATCCGCAACATCACAAACAGACCCAGTCCACAGATAGAGTGCATAGG 572
```

Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY 342 GAGCAGATAGCA---CAGGACCTGAGAAATGGCGCGCGAGGG-----GGCTTTGGCTCT 392
Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGln 176
QY 393 CGGGGTGACCCCGCAGAGCTCATCTGTGGCAGCGGGGCCCGACCGCAAGACGACGAA 452
Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY 453 GTGGACATCCCTTCGGCTCTCGGTGTCGCCACCCAGTATGCTGGTGGCTTATGGCAAG 512
Db 194 CysAspLeuProLeuArgLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
QY 513 GAGGGGGCCACCATCCGCAACATCACAAACAGACCCAGCTCCAGATAGACGTGCATAGG 572
Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY 573 AAGGAGACGCGAGTGCAGCTGAAAGAGCCATCAGTGTGCATCTCCACCTCGAGGGCTGC 632
Db 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
QY 633 TCCTCCGCTGTGAAGATGCTTGGAGATTATCATAAAGAGCTAAGGACACCAAAACG 692
Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
QY 693 GTGACGAGTTCCCTCGAAGATCTCGGCCATAATAACTTTTAGGGCGTCTCATGGC 752
Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY 753 AAGGAGGACGAGCACTGAGAGGTAGACAGATACGAGACAAATCACCATCTCC 812
Db 294 LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
QY 813 TCCTTCCAGACCTTACCTTTTACACCTCGAGGAGCACCATCTGTGAAGGGGGCCATC 872
Db 314 ProLeuGlnGluLeuThrLeuLysPheAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY 873 GAGAAATGTTGACGGCCGAGCAGGAGAAATAATGAAGAAATTCGGAGGCGCTATGAGAT 932
Db 334 GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
QY 933 GATGTGGCTCCATGAGC-----TCTCAGCTATCCTGGCTGACCTGAGCTGTGTA 986
Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeu 373
QY 987 GGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCTCCAGCAGCGTT 1037
Db 374 GlyLeuPheProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
QY 1038 ACTGGGGCTGCTCCTATAGCTCCTTTATGACGGTCCCGAGCAGGATGTTGCGAGGTG 1097
Db 394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
QY 1098 TTTATCCCGCCAGCAGTGGCGCCATCATCGGCAAGAGGGGCGAGCAGCATCAACAG 1157
Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlnHisLysGln 430
QY 1158 CTCTCCCGTTTCCAGCGCTCCATCAAGATTGCACACCCGAAACACCTGACTCCAAA 1217
Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY 1218 GTTCGTATGTTATCATCTAGCAGCGCCAGAGCCCAATTCAGGCTCAGGGAAGATC 1277
Db 451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY 1278 TATGGCAAACTCAAGGAGGAACTCTTTGGTCCCAAGCAGGAGTCTGAGCTGGAGCC 1337
Db 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY 1338 CATCATGCTGTCAGCAGTGGCGGGGTCTATTGGCAAGAGTGGAAAAACGGTG 1397
Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510

QY 1398 AACGACTTCAGAAATTTGACGCGAGCTGAGTGGTAGTACCAAGACAGACCCCTGAT 1457
Db 511 AsnGluLeuGlnAsnLeuSerAlaGluValValProArgAspGlnThrProAsp 530
QY 1458 GAGACGACGAGTCTATCGTGAATATCATCGGACATTTCTATGCCAGTCTAGATGGCTCAA 1517
Db 531 GluAsnAspGlnValValVallysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
QY 1518 CGGAAGATCCGAGACATCTCGGCCCGAGGTTAAGCAG--CAGCATCAGAAGGGA---CAG 1571
Db 551 ArgLysIleGlnIleLeuThrGlnValLysGlnHisGlnGlnLysAlaLeuGln 570
QY 1572 AGTAACACGCGCCAGCAGCAGCAGGAGAG 1598
Db 571 SerGlyProProGlnSerArgArgLys 579
RESULT 4
US-09-850-716A-348
; Sequence 348, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-850-716A-348
Alignment Scores:
Pred. No.: 1,52e-133 Length: 579
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best Local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 9 Gaps: 9
US-09-270-437D-7 (1-1946) x US-09-850-716A-348 (1-579)
QY 102 CGGTTAGGAGCGCGGAAATTCAAATCCGAATATTCACCCAGCTCCGATGGGAAGTA 161
Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluVal 96
QY 162 CTGGACAGCTGTGTGGCTCAGTATGTACAGTAGAAGTGTGAGCAAGTGAACACCGAG 221
Db 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AFTGACAGCGCATGTGTGAATGTCACTATTCCAAACCGGAGCAGACCGGCAAGCCATC 281
Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATGAAGCTGAATGCGCCACAGTGGAGAACCATGCCCTGAAGTCTCTACATCCCGAT 341
Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY 342 GAGCAGATAGCA---CAGGACCTGAGAAATGGCGCGCGAGGG-----GGCTTTGGCTCT 392
Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGln 176
QY 393 CGGGGTGACCCCGCAGAGCTCATCTGTGGCAGCGGGGCCCGACCGCAAGACGACGAA 452
Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY 453 GTGGACATCCCTTCGGCTCTCGGTGTCGCCACCCAGTATGCTGGTGGCTTATGGCAAG 512

QY 228 ACSGAGTGTGTAATGTCACCTATTCACCGGGAGCAGACCGAGGCAAGCCATCATGAG 287
 Db 119 ThrAlaValValAsnValThrTySerAsnArgGluGlnThrArgGlnAlaIleMetLys 138
 QY 288 CTGAATGGCCACAGTTGGAGAACCATGCGCTGAAGGTCTCTACATCCCGCATGAGCAG 347
 Db 139 LeuAsnGlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyrlleProAspGluGln 158
 QY 348 ATAGCACAGGAGCCTCGAAGATGGCGCGGAGGGGGCTTTGGCTCTCGGGGTTCAGCCCGC 407
 Db 159 IleThrGlnGlyProGluAsnGlyArgGlyGlyPheGlySerArgGlyGlnProArg 178
 QY 408 CAGGCTCACTGTGGCAGCGGGGGCCCCAGCCAGCAGCAGCAGCAAGTGGACATCCCTT 467
 Db 179 GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnProValAspIleProLeu 198
 QY 468 CGGCTCCTGTGCCACCCAGTATGGTGGTCCATATTGGCAAGGAGGGGGCCACCATC 527
 Db 199 ArgLeuLeuValProThrGlnTyrlleValGlyAlaIleleGlyLysGluGlyAlaThrIle 218
 QY 528 CGCAACATCAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAAGGAGAACCGAGGT 587
 Db 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisA-gLysGluAsnAlaGly 238
 QY 588 GCAGCTGAAGAACCATCAGTGTGACTCCACCCCTGAGGCTGCTCTCGCTTGTAAAG 647
 Db 239 AlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCysSerSerAlaCysLys 258
 QY 648 ATGATCTTGGAGATTATGCAAAAGAGGCTAAGAGACCAAAACCGCTGACGAGTCCC 707
 Db 259 MetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThrAlaAspGluValPro 278
 QY 708 CTGAAGATCTCGGCCCAATAAATCTTTAGGGCTCTCATTTGCAAGGAAGGACGGAAC 767
 Db 279 LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn 298
 QY 768 CTGAAGAGGTAGACAGATACCGAGCAAAATTCACCTCTCTCTGTTCCAGACCTT 827
 Db 299 LeuLysLysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeu 318
 QY 828 ACCCTTACACCCCTGAGAGACCATCACTGTGAAGGGGGCCATCGAGAAATTTGTGCAGG 887
 Db 319 ThrLeuTyAsnProGluArgThrIleThrValLysGlyAlaIleGluAsnCysCysArg 338
 QY 888 GCCGAGCAGGAATAATAAGAAAGTTCGGAGGCTATGAGAATGATGTGCTGCCATG 947
 Db 339 AlaGluGlnGluIleMetLysLysValArgGluAlaTyArgAsnAspValAlaAlaMet 358
 QY 948 AGC-----TCTACCTGATCCGCTGCTGAACTGGCTGTGTAGTCTTTTCCAGCT 1001
 Db 359 SerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaValGlyLeuPheProAla 378
 QY 1002 TCATCCAGCGAGTCCCGCCCTCCAGCAGCGTTACTGGGGTGTCTCCCTATAGTCC 1061
 Db 379 SerSerAlaValProProProSerSerValThrGlyAlaAlaProTySerSer 398
 QY 1062 TTTATGAGGCTCCCGAGCAGAGATGCTGAGGTGTTTATCCCGCCCGCAGGCAAGTGGGC 1121
 Db 399 PheMetGlnAlaProGluGlnGluMetValGlnValPheIleProAlaGlnAlaValGly 418
 QY 1122 GCCATCATCGGCAAGAGGGGAGCAGACATCAAAAGCTCTCCCGTTTCCAGCGCCCTCC 1181
 Db 419 AlaIleIleGlyLysLysGlyGlnHisIleLysGlnLeuSerArgPheAlaSerAlaSer 438
 QY 1182 ATCAAGATTGACACCCGAAACCTGACTCCAAAGTTCGTATGTTATCATCACTCGA 1241
 Db 439 IleLysIleAlaProProGluThrProAspSerLysValArgMetValValIleThrGly 458
 QY 1242 CCGCCAGAGGCCAAATCAAGGCTCAGGGAAGATCTATGGCAAACTCAAGGAGAGAAC 1301
 Db 459 ProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrlleGlyLysLeuLysGluLysAsn 478
 QY 1302 TTCCTTGGTCCCAAGAGGAAGTGAAGCTGGAGACCCACATACATCGTGTGCCAGCATCAGCA 1361

Db 479 PhePheGlyProLysGluGluValLysLeuGluThrHisIleArgValProAlaSerAla 498
 QY 1362 GCTGGCCGGGTTCATTGGCAAGGTGGAAAAACCGGTGAACGAGTTGCGAATTTGACGGCA 1421
 Db 499 AlaGlyArgValIleGlyLysGlyGlyLysThrValAsnGluLeuGlnAsnLeuThrAla 518
 QY 1422 GCTGAGTGTGTAGTACCAAGAGACACCCCTCATGAGAACCGAGGTTCATCGTGAAA 1481
 Db 519 AlaGluValValProArgAspGlnThrProAspGluAsnAspGlnValIleValLys 538
 QY 1482 ATCATCGACATTTCTATGCGAGTCAGATGGCTCAACCGAAGATCCGAGACATCTCGGCC 1541
 Db 539 IleIleGlyHisPheTyrlleAlaSerGlnMetAlaGlnArgLysIleArgAspIleLeuAla 558
 QY 1542 CAGTTTAAGCAGCAGCATCAAGAGGACAGAGTAAACCGGCCCGCAGGACCGAGGAAG 1598
 Db 559 GlnValLysGlnGlnHisGlnLysGlyGlnSerAsnLeuAlaGlnAlaArgArgLys 577

RESULT 3
 US-09-735-705-348
 ; Sequence 348, Application US/09735705
 ; Patent No. US20020052329A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Hosken, Nancy
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Li, Samuel X.
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: McNeill, Patricia D.
 ; APPLICANT: Fanger, Neil
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.455C14
 ; CURRENT APPLICATION NUMBER: US/09/735,705
 ; CURRENT FILING DATE: 2000-12-12
 ; NUMBER OF SEQ ID NOS: 419
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 348
 ; LENGTH: 579
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-735-705-348

Alignment Scores:
 Pred. No.: 1.52e-133 Length: 579
 Score: 1875.00 Matches: 380
 Percent Similarity: 84.68% Conservative: 51
 Best Local Similarity: 74.66% Mismatches: 62
 Query Match: 53.10% Indels: 16
 DB: 9 Gaps: 9

US-09-270-437D-7 (1-1946) x US-09-735-705-348 (1-579)

QY 102 CGGGTTAGGAGCGGAAAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTA 161
 Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluVal 96
 QY 162 CTGACAGCCTGCTGGCTCAGTATGGTACGTAGAGAACTGTGAGCACTGTGAGCAAGTGAACCCGAG 221
 Db 97 LeuAspSerLeuLeuValGlnTyrlleGlyValValGluSerCysGluGlnValAsnThrAsp 116
 QY 222 AGTGAGCGCAGTGGTGAATGTACCTATTCCACCGGAGCAGACACCGCAAGCCCATC 281
 Db 117 SerGluThrAlaValValAsnValThrTyrlleSerLysAspGlnAlaArgGlnAlaLeu 136
 QY 282 ATGAAGCTGAATGGCCACCATGTTGGAGAACCATGCGCTGAAGGTCTCTACATCCCGCAT 341

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 23, 2004, 11:15:41 ; Search time 92.8527 Seconds
(without alignments)
13126.119 Million cell updates/sec

Title: US-09-270-437D-7
Perfect score: 3531
Sequence: 1 gctgtagcggaggctggg.....atttcctcaggtttaaaa 1946

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1288442 seqs, 313154207 residues

Total number of hits satisfying chosen parameters: 2576884

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+ n2p.model -DEV=xl
-Q=/cgn2_1/USPTO_spool_p/US09270437/runat_23072004_095742_27560/app_query.fasta_1.5582
-DB=Published_Applications_AA -QPM=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITIS=bits -START=1 -END=-1 -MAIRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCOR=pct -THR MAX=100
-MAXLEN=2000000000 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp:*
- 3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pcp:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp:*
- 6: /cgn2_6/ptodata/2/pubpaa/PTUS_PUBCOMB.pcp:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09D_PUBCOMB.pcp:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pcp:*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pcp:*
- 18: /cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
------------	-------	-------	--------	-------	-------------

ALIGNMENTS

RESULT 1

US-10-313-986-500
; Sequence 500, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C19
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 500
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-313-986-500

Alignment Scores:

1	2509	71.1	577	15	US-10-313-986-500	Sequence 500, Appli
2	2491	70.5	577	9	US-09-873-637-2	Sequence 2, Appli
3	1875	53.1	579	9	US-09-735-705-348	Sequence 348, App
4	1875	53.1	579	9	US-09-850-716A-348	Sequence 348, App
5	1875	53.1	579	9	US-09-897-778-348	Sequence 348, App
6	1875	53.1	579	9	US-09-897-778-446	Sequence 446, App
7	1875	53.1	579	9	US-09-897-778-449	Sequence 449, App
8	1875	53.1	579	12	US-10-007-700-348	Sequence 348, App
9	1875	53.1	579	12	US-10-007-700-446	Sequence 446, App
10	1875	53.1	579	12	US-10-007-700-449	Sequence 449, App
11	1875	53.1	579	14	US-10-117-982-348	Sequence 348, App
12	1875	53.1	579	14	US-10-117-982-446	Sequence 446, App
13	1875	53.1	579	14	US-10-117-982-449	Sequence 449, App
14	1875	53.1	579	14	US-10-117-982-480	Sequence 480, App
15	1875	53.1	579	15	US-10-313-986-348	Sequence 348, App
16	1875	53.1	579	15	US-10-313-986-446	Sequence 446, App
17	1875	53.1	579	15	US-10-313-986-449	Sequence 449, App
18	1875	53.1	579	15	US-10-313-986-480	Sequence 480, App
19	1875	53.1	586	9	US-09-850-716A-427	Sequence 427, App
20	1875	53.1	586	9	US-09-897-778-427	Sequence 427, App
21	1875	53.1	586	12	US-10-007-700-427	Sequence 427, App
22	1875	53.1	586	14	US-10-117-982-427	Sequence 427, App
23	1875	53.1	586	15	US-10-313-986-427	Sequence 427, App
24	1875	53.1	589	15	US-10-313-986-486	Sequence 486, App
25	1873	53.0	579	9	US-09-735-705-176	Sequence 176, App
26	1873	53.0	579	9	US-09-850-716A-176	Sequence 176, App
27	1873	53.0	579	9	US-09-897-778-176	Sequence 176, App
28	1873	53.0	579	10	US-09-466-396A-176	Sequence 176, App
29	1873	53.0	579	12	US-10-007-700-176	Sequence 176, App
30	1873	53.0	579	14	US-10-117-982-176	Sequence 176, App
31	1873	53.0	579	15	US-10-313-986-176	Sequence 176, App
32	1871	53.0	579	14	US-10-117-982-484	Sequence 484, App
33	1871	53.0	579	15	US-10-313-986-484	Sequence 484, App
34	1763.5	49.9	620	9	US-09-764-864-1116	Sequence 1116, Ap
35	1757.5	49.8	587	15	US-10-313-986-501	Sequence 501, App
36	1697	48.1	556	14	US-10-097-340-147	Sequence 147, App
37	1697	48.1	556	15	US-10-648-593-182	Sequence 182, App
38	1573.5	44.6	555	15	US-10-262-445-40	Sequence 40, Appl
39	1560	44.2	422	16	US-10-408-765A-2088	Sequence 2088, Ap
40	1258	34.6	261	9	US-09-764-864-1114	Sequence 1114, Ap
41	1228	34.8	250	9	US-09-764-864-1532	Sequence 1532, Ap
42	653	18.5	171	9	US-09-764-864-1119	Sequence 1119, Ap
43	510	14.4	192	9	US-09-764-864-1117	Sequence 1117, Ap
44	502	14.2	171	9	US-09-764-864-1536	Sequence 1536, Ap
45	361	10.2	81	14	US-10-117-982-476	Sequence 476, App

This Page Blank (uspto)

QY	699	---GAGGTCCCTGAAGATCTCGCCCATATAACTTTGTAGGCGTCTCATTTGGCAAG	755
Db	184	GlyAsnAlaValGlnGluLeuMetIleProAlaSerLysAlaGlyLeuValIleGlyLys	203
QY	756	GAAGACGGAACCTGAAGAGGTAGACAGATACCGAGACAAATAATCACCATCTCTCCG	815
Db	204	GlyGlyGluThrIleLysGlnLeuGlnGluValGlyValLysMetValMet	221
QY	816	TTGCAAGACCTTACCTTTTACAACTTCAGAGGACC	860
Db	222	IleGlnAsp	237
QY	861	AAGGGGCGCATCGAATTTGTTCAGGCGCGAGAGAAATAATGAAGAAATTCGGGAG	920
Db	238	ThrGlyAspProTyLysValGlnGlnAlaLysGluMetValLeuGluLeuIleArgAsp	257
QY	921	---GCCTATGAGATGATGGCTGCCATGAGCTCTCACCTGATCCCT	965
Db	258	GlnGlyGlyPheArgGluValArgAsnGluTyGly	269
QY	966	GGCTGAACCTGGCTGTAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCGCT	1025
Db	269	---SerArgIleGlyGly	276
QY	1026	CCAGCAGCGTTACTGGGCTGCTCCTATAGCTCTTTTANGCGCTCCCGAGCAGGAG	1085
Db	270	---SerArgIleGlyGly	276
QY	1086	ATGGTGCAGGTGTTATCCCGCCCGCAGCAGTGGCGCCATCATCGCAAGAGGGGCGAG	1145
Db	277	GlyIleAspValProIleProArgPheAlaValGlyIleValIleGlyArgAsnGlyGlu	296
QY	1146	CACATCAACAGCTCTCCCGGTTTCGACGCGCTCCATCAAGATTGCACCCCGAA	1202
Db	297	MetIleLysIleGlnAsnAspAlaGlyValArgIleGlnPheLysProAspGly	316
QY	1203	---ACACCTGACTCCAAAGTTTCGTATGTTATCATCATCGACCGCCAGAG---GCCCAA	1256
Db	317	ThrThrProGlu	333
QY	1257	TTCAAGGCTCAG	1268
Db	334	HisAlaAlaGluIleIleThrAspLeuLeuArgSerValGlnAlaGlyAsnProGlyGly	353
QY	1269	---GGAGAAATCTATCGCAATCTATCGCAATCTATCGCAATCTATCGCAATCTATCGCAAT	1316
Db	354	ProGlyProGlyGlyArgGlyArgGlyArgGlyGlnGlyAsnTrpAsnMetGlyProPro	373
QY	1317	GAGGAAGTGAAGTGGACCCACATACGTGTCCAGCATCAGCAGCTGGCGGGTCAAT	1376
Db	374	GlyGlyLeuGln---GluPheAsnPheIleValProThrGlyLysThrGlyLeuIleIle	392
QY	1377	GCCTAAGTGGAAACCGTGAACGATTGTCAGATTTACCGCAGCTGAGGTGGTAGTA	1436
Db	393	GlyLysGlyGlyGluThrIleLysSerIleSerGlnGlnSerGlyAlaArgIleGluLeu	412
QY	1437	CCAAGAGACACGACCCCTGATGAGAACACCCAGGTC---ATCGTGAATAATCATCGGACAT	1493
Db	413	GlnArgAsnProProAsnAlaAspProAsnMetLysLeuPheThrIleArgGlyThr	432
QY	1494	---TTCTATCCAGTCCAGTGGCTCAACGGAAGAT	1534
Db	433	ProGlnGlnIleAspPyrAlaArgGlnLeuIleGluGluLysIleGlyGlyProValAsn	452
QY	1535	CCT---GGCCAGGTTAAGCAGCAGCATCAGAAAGGACAGAGTAACAGGCCCGCAGCAG	1591
Db	453	ProLeuGlyProProValProHisGlyProHisGlyVal	470
QY	1592	GA---GGAAGTGACAGCCCTC	1624
Db	470	IyProProGlyProProGlyProGlyThrProMetGlyProTyLysProAlaProTyra	490

Search completed: July 23, 2004, 11:19:34
Job time : 44.4913 secs

```
Db 413 GlnArgAsnProProAsnAlaAspProAsnMetLysLeuPheThrIleArgGlyThr 432
QY 1494 -----TTCATGCCAGTCAGATGGCTCAACGGAAGAT-----CCGAGACAT 1534
Db 433 ProGlnGlnIleAspTyrAlaArgGlnLeuIleGlnLysIleGlyGlyProValAsn 452
QY 1535 CCT---GGCCCGAGTTAAGCAGCAGCATCAGAGGGACAGAGTAACAGGCCCGCAGGCACG 1591
Db 453 ProLeuGlyProValProHisGlyProHisGlyVal-----ProGlyPro-HisG 470
QY 1592 GA-----GGAGTCACCGCCCTC-----CCTGTCCCTTNGA 1624
Db 470 lYProGlyProProGlyProGlyThrProMetGlyProTyrAsnProAlaProTyrA 490
QY 1625 GTCAGGA-----1632
Db 490 snProGlyProProGlyProAlaProHisGlyProProAlaProTyrAlaProGlnGlyT 510
QY 1633 -----CAACACGGGCGAATCGAGAGTGTCTCCCGG 1669
Db 510 rpGlyAsnAlaTyrProHisTrpGlnGlnAlaProProAsp-----ProA 526
QY 1670 CAGGCGTCGAGATCAGTGGGAATCCGGACACNTGGCGGGCTGTAGATCAGGTTTGGC 1729
Db 526 lAlysAlaGlyThrAspProAsnSerAla-AlaTrpAlaAlaTyr-----TyrAla 542
QY 1730 CACTTGATTGAGAAAGATGTTCCAGTCAGGAGAACCTGATCTNTCAGCGCCCAACACCCAC 1789
Db 543 HisTyrTyrGlnGlnAla-----GlnProProProAlaAla 555
QY 1790 CCAATGGC---CCAACACTGTNTGCCCTCCGGGTGTCAAGAAATNTAGCGAAGCCAC 1846
Db 556 ProAlaGlyAlaProThrThrThrGlnThrAsnGlyGlnGlyAsp-----570
QY 1847 TTTTAAACGTGGATTGTTTAAAGAAAGCTCTCCAGGCGCCCAAGAGGGTGGATCACACC 1906
Db 571 -----GlnGlnAsnProAlaProAlaGlyGlnValAspTyrThr 583
QY 1907 -----TCAGTGGGAAGAAAAATAAAATTTCTCTCAGGT 1939
Db 584 LysAlaTrpGluGluTyrTyrLysLysMetGlyGlnAlaValProAlaProThrGly 602

RESULT 15
PCT-US94-01782-2
; Sequence 2, Application PC/TUS9401782
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES
; APPLICANT: AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF
; APPLICANT: HEALTH AND HUMAN SERVICES
; TITLE OF INVENTION: NOVEL FUSE BINDING
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01782
; FILING DATE: 22-FEB-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/021,608
; FILING DATE: 22-FEB-1993
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: WILLIAM S. FEILER
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4063PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 644
; TYPE: Amino Acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; MOLECULE TYPE: Peptide/Protein
; HYPOTHETICAL: No
; ORIGINAL SOURCE:
; ORGANISM: Human
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE: HL60
; ORGANELLE:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: Amino Acid 149 (Xaa) is Met or Ile
; PCT-US94-01782-2

Alignment Scores:
Pred. No.: 1.83e-13 Length: 644
Score: 245.00 Matches: 148
Percent Similarity: 36.46% Conservative: 93
Best Local Similarity: 22.39% Mismatches: 219
Query Match: 6.94% Indels: 202
DB: 5 Gaps: 33

US-09-270-437d-7 (1-1946) x PCT-US94-01782-2 (1-644)
QY 309 AACATGCTCCCTGAAGGTCTCTACATCCCGAGTACAGATACACAG-----GGACCT 362
Db 28 AsnAspAlaPheLysAspAlaLeuGlnArgAlaArgGlnIleAlaAlaLysIleGly 47
QY 363 GAGATGGG-----CGCCGAGGGGGTCTCGGGTCTCGGGTCAAGCCCGCAGGGC 413
Db 48 AspAlaGlyThrSerLeuAsnSerAsnAspTyrGlyGlyGlnLysArg-----65
QY 414 TCACCTGTGGCAGCGGG-----431
Db 66 ---ProLeuGluAspGlyAspGlnProAspAlaLysLysValAlaProGlnAsnAspSer 84
QY 432 -----GCCCGAGCCAGCAGCAGTGGACATCCCTCT-----CGG 470
Db 85 PheGlyThrGlnLeuProProMetHisGlnGlnGlnSerArgSerValMetThrGlu 104
QY 471 CTCCTGTGTGCCACCCAGTATGGGTGCCATTATTGGCAAGGAGGGGGCCACATCCGC 530
Db 105 TyrLysValProAspGlyMetValGlyPheIleIleGlyArgGlyGlyGlnIleSer 124
QY 531 AACATCACAACAGACCCAGTCCAGATAGACGTAGACGATAGGAGGAGAACCGAGTGCA 590
Db 125 ArgIleGlnGlnLysSerGlyCysLysIleGlnIle---AlaProAspSerGlyGlyLeu 143
QY 591 GCTGAAAAGCCATCAGTGTGCACTCCACCTGAGGGGTGCTCTCCGCTTTAAGATG 650
Db 144 ProGluArgSerCys***LeuThrGlyThrProGluSerValGlnSerAlaLysArgLeu 163
QY 651 ATCTTGGAGATTATGCATAAA-----GAGGCTTAAGACACCAACACGGGTGAC-----698
Db 164 LeuAspGlnIleValGluLysGlyArgProAlaProGlyPheHisIleGlyAspGlyPro 183
```


Db 386 LysThrGlyLeullelleGlyLysGlyGlyGluThrIleYsserIleSerGlnGlnSer 405
Qy 1419 GCAGCTGAGGTGTAGTACCAAGACAGACCCCTGATGAGACACAGGTC---ATC 1475
Db 406 GlyAlaArgileGluGlnArgAsnProProGlnAlaAspProAsnMetLysLeu 425
Qy 1476 GTGAAATCATCGACAT-----TTCTATGCCAGTCCAGTCCCAAGGAG 1523
Db 426 PheThrIleArgGlyThrProGlnGlnIleAspTyrAlaArgGlnLeulleGluGluLys 445
Qy 1524 AT-----CGGACATCCT---GGCCCGTGAAGCAGCAGCATCAGAGGACACAG 1573
Db 446 IleGlyGlyProValAsnProLeuGlyProProValProHisGlyProHisGlyVal--- 464
Qy 1574 TAACAGCCGCCAGCAGGA-----GGAAGTGACAGCAGCCCTC----- 1611
Db 465 ---ProGlyPro--HisGlyProProGlyProProGlyProGlyThrProMetGlyProT 483
Qy 1612 ---CCTGCTCCCTTNGAGTCCAGGA----- 1632
Db 483 yrAsnProAlaProTyrAsnProGlyProProGlyProAlaProHisGlyProProAlaP 503
Qy 1633 -----CAACAAGCGGCAGAAATCG 1651
Db 503 roTyrAlaProGlnGlyTyrGlyAsnAlaTyrProHisTyrGlnGlnGlnAlaProProA 523
Qy 1652 AGAGTGTGCTCTCCCGCAGCGCTGAGAATGAGTGGGAATCCGGGACACNTGGCGCGG 1711
Db 523 sp-----ProAlaLysAlaGlyThrAspProAsnSerAla-AlaTyrAlaAla 538
Qy 1712 CTGTAGATCAGTTTGCCCATCTGATTCGAGAAGATGTCAGTGAGGAACCTGATCTN 1771
Db 539 Tyr-----TyrAlaHisTyrTyrGlnGlnGlnAla----- 548
Qy 1772 TCAGCCCCAACACCCCAATGGC---CCAACACTGTNTGCCCTCGGSGTGTCCAGA 1828
Db 549 GlnProProAlaAlaProAlaGlyAlaProThrThrThrGlnThrAsnGlyGlnGly 568
Qy 1829 AATTNTAGCGCAGGCACTTTTAAACGTGGATTGTTTAAAGAAGCTCTCCAGGCCCCACC 1888
Db 569 Asp-----GlnGlnAsnProAlaProAla 576
Qy 1889 AAGAGGTGTATCATCACC-----TCAGTGGGAAGAAATA 1924
Db 577 GlyGlnValAspTyrThrLysAlaTrpGluGluTyrTyrLysMetGlyGlnAlaVal 596
Qy 1925 AAATTTCCTTCAGT 1939
Db 597 ProAlaProThrGly 601

RESULT 12
US-09-261-855-18
; Sequence 18, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-18
Alignment Scores:
Pred. No.: 4,67e-14 Length: 49
Score: 245.00 Matches: 49
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.94% Indels: 0
DB: 3 Gaps: 0
US-09-270-437D-7 (1-1946) x US-09-261-855-18 (1-49)
Qy 714 ATCCCTGCCCATATAACTTTGTAGGGCTCTCATTTGGCAAGGACGGAACCTGAAG 773
Db 1 IleLeuAlaHisAsnAsnPheValGlyArgLeulleGlyLysGluGlyArgAsnLeuLys 20
Qy 774 AAGCTAGACGAGATACCGAGACAAAATACACCATCTCTCTCGTTGCAAGACCTTACCCTT 833
Db 21 LysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeuThrLeu 40
Qy 834 TACAACCTGTAGAGACCATCAGTGTG 860
Db 41 TyrAsnProGluArgThrIleThrVal 49

RESULT 13
US-08-021-608D-2
; Sequence 2, Application US/08021608D
; Patent No. 5580760
; GENERAL INFORMATION:
; APPLICANT: LEVENS, DAVID L., DUNCAN,
; APPLICANT: ROBERT C., AND AVIGAN, MARK I.
; TITLE OF INVENTION: NOVEL FUSE BINDING
; TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/021,608D
; FILING DATE: 22-FEB-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAM S. FEILER
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4063
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 644
; TYPE: Amino Acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; MOLECULE TYPE: Peptide/Protein
; HYPOTHETICAL: No
; ORIGINAL SOURCE:
; ORGANISM: Human
; CELL LINE: HL60
; FEATURE:
; OTHER INFORMATION: Amino Acid 149 (Xaa) is Met or Ile
US-08-021-608D-2
Alignment Scores:
Pred. No.: 1.83e-13 Length: 644

APPLICANT: THE GOVERNMENT OF THE UNITED STATES
APPLICANT: AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF
APPLICANT: HEALTH AND HUMAN SERVICES
TITLE OF INVENTION: NOVEL FUSE BINDING
TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:

; OTHER INFORMATION: Amino Acid 148 (Xaa) is Met or Ile
 ; OTHER INFORMATION: Amino Acid 148 (Xaa) is Met or Ile
 US-08-726-160-10

Alignment Scores:	8.81e-14	Length:	643
Pred. No.:	248.50	Matches:	140
Score:	37.80%	Conservative:	97
Percent Similarity:	23.33%	Mismatches:	202
Best Local Similarity:	7.04%	Indels:	189
Query Match:	1	Gaps:	32
DB:			

US-09-270-437D-7 (1-1946) x US-08-726-160-10 (1-643)

336	Qy	CCCGAT---	GAGCAGATAGACAGGACCTGTGAGAAATGGGCGCGAGGGGCTTTGGCTCT	392
			:::::	
73	Db	ProAspAlaIlysluysValAla----	ProGlnAsn-----AspSerPheGlyThr	87
393	Qy	CGGGGTGAGCCCGCCAGGGGCTCACCTGTGGCAGCGGGGCCCCAGCAAGCAGACAGCAAA	452	
		:::		
88	Db	GlnLeuProProMetHis-----	GlnGlnGlnArg	97
453	Qy	GTGGACATCCCCCTTCGGCTCCTGGTGCCACCACCACTATGTGGGTGCCATTATTGGCAAG	512	
		:::		
98	Db	SerValMetThrGluGluTuTyLysValProAspGlyMetValGlyPheIleIleGlyArg	117	
513	Qy	GAGGGGCCACCATCCGCAACATCACAAACAGACCCAGTCCCAAGATAGACGTGCATAGG	572	
118	Db	GlyGlyGluGlnIleSerArgIleGlnGlnGluSerGlyCysIleIleGlnIle----	Ala	136
573	Qy	AAGGAGAACCGAGGTGCAGCTGAAAGACCATCAGTGTGCACCTCCACCCCTAGGGGTGC	632	
137	Db	ProAspSerGlyGlyLeuProGluArgSerCys**LeuThrGlyThrProGluSerVal	156	
633	Qy	TCCTCCGCTTGTAAAGATGATCTTGGAGATTATGCATAAA-----	GAGGCTAAGGACACC	686
157	Db	GlnSerAlaIysArgLeuLeuAspGlnIleValGlnIlysluysGlyArgProAlaProGlyPhe	176	
687	Qy	AAAAACGGCTCAC-----	GAGGTTCCCTGAAGATCTCGGCCCATATAATCTTGTA	737
177	Db	HisGlyAspGlyProGlyAsnAlaValGlnGluIleMetIleProAlaSerLysAla	196	
738	Qy	GGGGCTTCTATTCGCAAGGAGGACCGAACTCGAAGAAAGGTAGACGAAGATACCGAGACA	797	
197	Db	GlyLeuValIleGlyLysGlyGlyGluThrIleLysGlnLeuGlnGluArgAlaGlyVal	216	
798	Qy	AAATCACCATCTCCTCGTTTGCAGACCTTTACCCCTTTACAACCTGAGAGACC-----	851	
217	Db	LysMetValMet-----	IleGlnAsp-----GlyProGlnAsnThrGlyAla	230
852	Qy	-----ATCACTGTGAAGGGGGCCATCGAGAATTGTTCAGGGCGCGCAGGAAATA	902	
231	Db	AspLysProLeuArgIleThrGlyAspProTyrIysValGlnGlnAlaLysGluMetVal	250	
903	Qy	ATGAGAAAGTTCGGAG-----	GCCTATGAGATGATGGCTGCCCATG	947
251	Db	LeuGluIleArgAspGlnGlyPheArgGluValArgAsnGluTyrGly-----	268	
948	Qy	AGCTCTCACCTGATCCCTGGCCTGAACCTGGCTGCTAGTCTTTTCCAGGCTTCATCC	1007	
268	Db	-----	-----	268
1008	Qy	AGCGCATCCCGCCGCTCCAGCAGGTTACTGGGGCTGCTCCCTATAGTCTCTTTATG	1067	
269	Db	-----SerArgIleGlyGly-----	-----	273
1068	Qy	CAGGCTCCGAGCAGGAGATGGTCCAGTGGTTTATCCCGCCAGGCGAGTGGGGCCCATC	1127	
274	Db	-----AsnGluGlyIleAspValProIleProArgPheAlaValGlyIleVal	289	
1128	Qy	ATCGGCAAGAGGGCGCAGCATCAACACAGCTCTCCCGTTGTCGACGGCCTCCATCAAG	1187	
290	Db	IleGlyArgAsnGlyGluMetIleLysLysIleGlnAsnAspAlaGlyValArgIleGln	309	

Qy	1188	ATTGCAACCCGAA-----ACACTGACTCAAAAGTTTCGTATGGTTATCATCATCGA	1241
Db	310	PheLysProAspGlyThrThrProGlu-----ArgileAlaGlnileThrGly	326
Qy	1242	CCGCCAGAG---GCCCAATTCAAGGCTCAG-----	1268
Db	327	ProProAspArgCysGlnHisalaAlaGluIlelleThrAspLeuLeuArgSerValGln	346
Qy	1269	-----GGAAGAATCTATGGCAAACTATGCGAACTCAAGGAGGAG	1298
Db	347	AlaGlyAsnProGlyGlyProGlyProGlyGlyArgGlyArgGlyGlyGlnGlyAsn	366
Qy	1299	AACTCTTTGGTCCCAAGAGGAAAGTGAAGCTGGAGCCCATACGTGTGCCAGCATCA	1358
Db	367	TrpAsnMetGlyProProGlyGlyLeuGln---GluPheAsnPhelleValProThrGly	385
Qy	1359	GCAGCTGGCCGGTCAATTGGCAAAAGTGTGAAAACGGTCAACGAGTTGCAGAATTTGACG	1418
Db	386	LysThrGlyLeuIlelleGlyGlyGlyGlyGluThrIleLysSerIleSerGlnGlnSer	405
Qy	1419	GCAGCTGAGTGGTAGTACCAAGAGACCAGACCCCTGATGAGAACACCCAGGTC---	1475
Db	406	GlyAlaArgileGluLeuGlnArgAsnProProProAsnAlaAspProAsnMetLysLeu	425
Qy	1476	GTGAATCATCCGACAT-----TTCTATGCCAGTCAGTCAGTGGCTCAACGGAG	1523
Db	426	PheThrIleArgGlyThrProGlnGlnIleAspTyrAlaargGlnleuGluGlnLys	445
Qy	1524	AT-----CCGAGACATCCT---GGCCAGGTTTAAGCAGCAGCATCAGAAGGACAGAG	1573
Db	446	IleGlyGlyProValAsnProLeuGlyProProValProHisGlyProHisGlyVal---	464
Qy	1574	TAAACAGGCCAGCGCGCGGA-----GGAAAGTGACACGACCCCTC-----	1611
Db	465	---ProGlyPro---HisGlyProProGlyProProGlyProGlyThrProMetGlyProT	483
Qy	1612	-----CCTGTCCTTNGAGTCCAGGA-----	1632
Db	483	YrAsnProAlaProTyrAsnProGlyProProGlyProAlaProHisGlyProProAlaP	503
Qy	1633	-----CAACAACGGCGAGAAATCG	1651
Db	503	roTyrAlaProGlnGlyTyrGlyAsnAlaTyrProHisIleProGlnGlnAlaProProA	523
Qy	1652	AGAGTGTGCTCCCGCGCGGCGCTAGAAATGATGGGAATCGGGACACNTGGGCGGG	1711
Db	523	sp-----ProAlaLysAlaGlyThrAspProAsnSerAla-AlaTrpAlaAla	538
Qy	1712	CTGTAGATCAGGTTTGCCCATCTTGATTGAGAAAGATGTTCCAGTGAAGAACCTGTCTN	1771
Db	539	Tyr-----TyrAlaHisTyrTyrGlnGlnAla-----	548
Qy	1772	TCAGCCCCAACACCCACCAATTTGCG---CCAACACTGTNTGCCCTCGGGGTGCAGA	1828
Db	549	GlnProProAlaAlaProAlaGlyAlaProThrThrGlnThrAsnGlyGlnGly	568
Qy	1829	AATTNTAGCCAGGCACITTTAAACGTGGATGTTTAAAGAGCTCTCCAGGCCCCACC	1888
Db	569	Asp-----GlnGlnAsnProAlaProAla	576
Qy	1889	AAGAGGGTGATCACACC-----TCAGTGGGAAGAAAAATA	1924
Db	577	GlyGlnValAspTyrThrLysAlaIleValGluGluTyrTyrLysLysMetGlyGlnAlaVal	596
Qy	1925	AAATTCCTTCAGGT	1939
Db	597	ProAlaProThrGly	601

RESULT 11
PCT-US94 -
; Sequenc
; GENERA


```
QY 813 TCCTGTCAGACCTTACCCCTTTACACCTGAGAGGACCATCATCTGTGAAGGGGGCCATC 872
Db 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY 873 GAGAAATTTGTCAGGGCGGACAGCAATATGAAGAAAGTTCCGGAGGCTATGAGAAT 932
Db 334 GluThrCysAlaLysAlaGluGluGluLeuMetLysLysIleArgGluSerTyrGluAsn 353
QY 933 GATGTGCTGCCATGAGC-----TCTCACCTGATCCCTGGCCCTGAACCTGTGCTGTGA 986
Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeu 373
QY 987 GGTCTTTTCCAGGCTTCATCCAGCGCAGTCCCGCG-----CTCCAGCAGCGTT 1037
Db 374 GlyLeuPheProThrSerGlyMetProProThrSerGlyProProThrSerAlaMet 393
QY 1038 ACTGGGCTGCTCCCTATAGCTCCCTTTATGAGGCTCCCGAGCAGGAGATGGTGCAGTG 1097
Db 394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisGln 410
QY 1098 TTATATCCCGCCAGGAGTGGGGCCCATCATCGGCAAGAGAGGGCGCGCATCAAAACAG 1157
Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGln 430
QY 1158 CTCTCCCGGTTTGGCAGCGCTCCATCAAGATTGCACCCCGCAACACCTGCTCCAAA 1217
Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY 1218 GTTGCTGTTGTTATCATCTGACCGCCAGGAGGCCCAATTCAGGCTCAGGGAAGATC 1277
Db 451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY 1278 TATGGCAACTCAAGGAGGAGAACTTTCTTGGTCCCAAGAGAGAGTGAAGCTGGAGAC 1337
Db 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY 1338 CACATAGCTGTGCAGCATCAGCAGCTGCGCGGTCAATTCGCAAGGTGGAAGAAACGGTG 1397
Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
QY 1398 AACGAGTTCAGAAATTTGACGGCAGCTGAGTGTGTAGTACCAAGACAGACCCCTGTAT 1457
Db 511 AsnGluLeuGlnAsnLeuSerAlaGluValValProArgAspGlnThrProAsp 530
QY 1458 GAGACGACCATCTCGTGAATCATCGGACATTCGACATTCATGCGAGTCAGTGGCTCAA 1517
Db 531 GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
QY 1518 CGAAGATCCGAGACATCTCGCCAGGTTAAGCAG---CAGCATCAGAGGGA---CAG 1571
Db 551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnLysAlaLeuGln 570
QY 1572 AGTAACAGGCCAGGCGACGAGGAG 1598
Db 571 SerGlyProProGlnSerArgLys 579
```

RESULT 8

```
US-09-606-421B-176
; Sequence 176, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasar A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
```

```
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRN
; ORGANISM: Homo sapiens
US-09-606-421B-176

Alignment Scores:
Pred. No.: 8,4e-161 Length: 579
Score: 1873.00 Matches: 380
Percent Similarity: 84.48% Conservative: 50
Best Local Similarity: 74.66% Mismatches: 63
Query Match: 53.04% Indels: 16
DB: 4 Gaps: 9
```

US-09-270-437D-7 (1-1946) x US-09-606-421B-176 (1-579)

```
QY 102 CGGTTAGAGAGCGGAAATTTCAAATCCGAATATTTCCACCCAGCTCCGATGGAAAGTA 161
Db 77 ArgGluArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTyrGluVal 96
QY 162 CTGGACAGCCTGCTGGCTCAGTATGTTACGTAGTAGAGAACTGTGAGCAAGTGAACACGAG 221
Db 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGAGAGCGCAGTGTGTAATGTCACCTATTTCCAAACGGGAGCAGACCCAGGCAATC 281
Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATGAAGCTGAATGCCACCCAGTTGGAGAACCATCGCTCAAGGTCTCTACATCCCGCAT 341
Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY 342 GAGCAGATAGCA---CAGGACCTGAGATGGCGCCGAGGG-----GGCTTTGGCTCT 392
Db 157 GluMetAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGln 176
QY 393 CGGGGTGAGCCCGCCAGGCTCACCTGTGGCAGCGGGGGCCCGCCAGCAGCAGCAACAA 452
Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY 453 GTGACATCCCTTCCGCTCCTGCTGCTGGTGGCCACCCAGTATGTGGGTGCCCATTTATGGCAAG 512
Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
QY 513 GAGGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCCAAGATAGACGTGCATAGG 572
Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY 573 AAGGAGACGCGAGGTGCAGTGAAGAGCCATCAGTGTGCACCTCCACCTCCAGGGGTGC 632
Db 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
QY 633 TCTTCGCTGTGAAGATGATCTTGGAGATTATGCATATAAGAGGCTAAGACACCAAAACG 692
Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
QY 693 GCTGACGAGGTTCCTCCCTGAGATCCCTGGGCCCATATAACTTTGTAGGGGTCTCANTGGC 752
Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY 753 AAGGAGGACGGAACCTGGAAGAGGTAGAGAGATACCGAGACACCAAAATACCATCTCC 812
Db 294 LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
QY 813 TCGTTCAAGACCTTACCTTTTACACCTTGAGAGGACCATCCTGTGTGAAGGGGGCCATC 872
Db 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
```

QY 693 GCTGACGAGGTTCCCTGAAGATCCTGGCCCATATAACTTTGTAGGCGGTCTCATTCGC 752
Db 274 ThrGluGluLeuProLeuLysLeuAlaHisAsnAsnProHisLeuGlnValGlyArgLeuLeuGly 293
QY 753 AAGGAGGACGACCTGAGGAGGTAGAGACATACCGAGACAAAATCACCATTCC 812
Db 294 LysGluGlyArgAsnLeuLysLeuGlnAspThrLysLeuLeuLeuLeuLeuLeuLeuLeuLeu 313
QY 813 TCGTTGCAAGACCTTACCTTTACAACTCAGAGGACCATCAGTGTGAAGGGGCCATC 872
Db 314 ProLeuGlnGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 333
QY 873 GAGAAATTTTTCAGGCGCGAGCAGGAAATAATGAGAAAGTTCCGAGGCGCTTATGAGAA 932
Db 334 GluThrCysAlaLysAlaGluGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 353
QY 933 GATGTGGTGCATGAGC-----TCTCACTGATCCCTGGCTGAACCTGCTCTGTA 986
Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 373
QY 987 GGTCTTTTCCAGCTTCATCAGCGCAGTCCCGCG-----CCTCCAGCAGCGTT 1037
Db 374 GlyLeuPheProProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
QY 1038 ACTGGGGTGTCTCCCTATAGCTCTTATGCAAGGCTCCCGAGCAGGAGATGTCAGGTG 1097
Db 394 Thr-----ProProTyProGlnPheGluGlnSer---GluThrGluThrValHisGln 410
QY 1098 TTTATCCCGCCGAGGAGTGGGCGCCATCATCGCGAAGAGGGCAGCAGCATCAACAG 1157
Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGln 430
QY 1158 CTCTCCCGGTTTGGCAGCGCTCCATCAAGATTGACACCCGAAACACCTCACTCCAAA 1217
Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY 1218 GTTCTATGTTTATCATCTGACCGCGCAGAGGCGCCCAATTCAGGCTCAGGAGAAATC 1277
Db 451 ValArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY 1278 TATGCGAACTCAAGGAGAACTTCTTTGTCCTCCAGGAGGAGTGAAGTGGAGACC 1337
Db 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY 1338 CACATAGGTGTCAGCATCAGCAGTGGCGCGGTCAATGGCAAGGTGGAAGAACGCTG 1397
Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
QY 1398 AACGAGTTGCAAGATTTGACGCGAGTGGTAGTACCAAGAGACCGAGCCCTGAT 1457
Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
QY 1458 GAGACGACGAGTCTCTGTAATCATCGACATTTCTATCCAGTCCAGTGGCTCA 1517
Db 531 GluAsnAspGlnValValLysIleThrGlyHisPheTyAlaCysGlnValAlaGln 550
QY 1518 CGAAGATCCGAGACATCTCTGGCCAGGTTAAGCAG---CAGCATCAGAGGGA---CAG 1571
Db 551 ArgLysIleGlnGluLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
QY 1572 AGTAACCGCGCAGCAGGAGGAG 1598
Db 571 SerGlyProGlnSerArgLys 579

RESULT 7

US-09-542-615A-176
; Sequence 176, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-542-615A-176
Alignment Scores:
Pred. No.: 8,4e-161 Length: 579
Score: 1873.00 Matches: 380
Percent Similarity: 84.48% Conservative: 50
Best Local Similarity: 74.66% Mismatches: 63
Query Match: 53.04% Indels: 16
DB: 4 Gaps: 9
US-09-270-437D-7 (1-1946) x US-09-542-615A-176 (1-579)

QY 102 CGGTTAGAGCGCGAAATTCAAATCCGAATATTCACCCAGCCTCGATGGGAAGTA 161
Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluVal 96
QY 162 CTGGACAGCGTCTGGCTCAGTATGTACAGTAGAGAACTGTGAGCAAGTGAACACCGAG 221
Db 97 LeuAspSerLeuLeuValGlnTyrglyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGAGCGGAGTGTGAATGTACCTATTCCAAACCGGAGCAGACCCAGCAAGCCATC 281
Db 117 SerGluThrAlaValValAsnValThrTyrSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATGAAGCTGAATGGCCACAGTTGGAGAACCATGCTGCAAGTCTCTACATCCCGAT 341
Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrlleProAsp 156
QY 342 GAGCAGATAGCA---CAGGACCTGAGAAATGGCGCGCAGGG---GGCTTTGGCTCT 392
Db 157 GluMetAlaAlaGlnGlnAsnProLeuGlnProArgGlyArgArgGlyLeuGlyGln 176
QY 393 CGGGGTTCAGCCCGCGCAGGCTCACCTGTGGCAGCGGGGGCCCGCCAGCAGCAAG 452
Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY 453 GTGGACATCCCGCTTCGGCTCTGCTGGTCCCGCCAGCATGTGGTCTCCATTTATGGCAG 512
Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
QY 513 GAGGGGCGCACCATCCGCAACATCAAAAACAGACCCAGCTCCCAAGATAGACGTGATAG 572
Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY 573 AAGGAGAACCGAGGTGCAGCTGAAAAGCCATCATGTCAGTGTGCATCCACCTGAGGGTGC 632
Db 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
QY 633 TCCTCGCTGTGAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCAAAACG 692
Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
QY 693 GCTGACGAGGTTCCCTCGAAGATCTCTGGCCCATATAACTTTGTAGGCGGTCTCATTCGC 752
Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuLeuGly 293
QY 753 AAGGAGGACGGAACCTGAGAGAGGTAGAGCAGAGATCCGAGACAAAATCACCATCTCC 812
Db 294 LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer 313

```
Db      214  GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY      573  AAGGAGAACGAGGTGAGCTGAAAGCCATCAGTGTGCATCTCACCCCTGAGGGCTGC 632
Db      234  LysGluAsnAlaGlyAlaAlaGlySerIleThrIleLeuSerThrProGluGlyThr 253
QY      633  TCCCTCCGCTTGAAGATGATCTCGAGATTATGATAAAGAGGCTAAGGACACCAACG 692
Db      254  SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIlePhe 273
QY      693  GCTGACGAGTCCCTCAAGATCCTGCCCATATACTTTGTAGGGCGTCTCATTCGC 752
Db      274  ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY      753  AAGGAAGGACGACCTGGAAGAGGTAGACAAATACCGAGACAAAATCACCATCTCC 812
Db      294  LysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
QY      813  TCGTGTCAAGACCTTACCTTTACAACCTGAGAGGACCATCTGTGTGAAGGGGCCATC 872
Db      314  ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY      932  GAGATTGTGCGAGGCGGACGAGAAATATGAAGAAAGTTCCGGAGGCGCTATGAGAT 932
Db      334  GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
QY      933  GATGTGCTGCCATGAGC-----TCTCACTGATCCTGGCCTGAACCTGGCTGTGTA 986
Db      354  AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeu 373
QY      987  GGTCTTTTCCAGCTTCATCCAGCGCATGCCGCG-----CTCCACGACGAGTT 1037
Db      374  GlyLeuPheProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
QY      1038  ACTGGGCTGCTCCTTAGCTCCTTTATGAGCTCCCGAGCAGGAGATGTTGACAGTG 1097
Db      394  Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisGln 410
QY      1098  TTTATCCCGCCGAGGAGTGGCGCCATCATCGGCAAGAGGGGCGACATCAACAG 1157
Db      411  PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyHisIleLysGln 430
QY      1158  CTCTCCGGTTGCGAGCGCTCCATCAAGATTGCACCAACCCGAAACCTGTACTCCAAA 1217
Db      431  LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY      1218  GTTCGTATGTTATCATCTGACCGCCAGAGCCCAATTCAGGCTCAGGGAAGATC 1277
Db      451  ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY      1278  TATGGCAACTCAAGGAGAGACTCTTTGGTCCCAAGGAGGAGTGAAGCTGGAGACC 1337
Db      471  TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY      1338  CACATACGTGTGCGACATCAGCAGCTGGCGGGTCTATTGGCAAGAGTGAAGGCGTG 1397
Db      491  HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
QY      1398  AACGAGTGCAGAAATTGACGCGAGCTGAGGTGGTAGTACCAAGACACGACCCCTGAT 1457
Db      511  AsnGluLeuGlnAsnLeuSerAlaGluValValProArgAspGlnThrProAsp 530
QY      1458  GAGAACGACCATCTATCGTGAATCATCGGACATTTCTATCCAGTACAGTGGCTCA 1517
Db      531  GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
QY      1518  CGAAGATCCGACATCTCTGGCCCGAGTTTAAGCAG---CAGCATCAGAAAGGA---CAG 1571
Db      551  ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
QY      1572  AGTAACGAGCCCGACGACGAGGAG 1598
Db      571  SerGlyProProGlnSerArgArgLys 579
```

```
RESULT 6
US-09-480-884A-176
; Sequence 176, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-480-884A-176

Alignment Scores:
Pred. No.:      8.4e-161      Length:      579
Score:          1873.00      Matches:     380
Percent Similarity: 84.48%      Conservative: 50
Best Local Similarity: 74.66%      Mismatches:  63
Query Match:      53.04%      Indels:      16
DB:               4          Gaps:         9
```

US-09-270-437d-7 (1-1946) x US-09-480-884A-176 (1-579)

```
QY      102  CGGGTTAGGAGCGGAAATTCAGAAATTCACCCAGCTCCGATCGGAAGTA 161
Db      77  ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProHisLeuGlnProGluVal 96
QY      162  CTGACAGCGCTGCTGCTCAGTATGTGTACAGTACAGAACTGTGAGCAAGTGAACACCGAG 221
Db      97  LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY      222  AGTGAGACGCGCAGTGTGATGTCACTATTCACCGGAGCAGACGAGGAGGAGCATC 281
Db      117  SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
QY      282  ATCAAGCTGAATGCCACCGAGTGGAGAACCATCGCTGAAGCTCTCTACATCCCGCAT 341
Db      137  AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY      342  GAGCAGATAGCA---CAGGGACCTGAGATGGCGCGCGAGGG-----GGCTTTGGCTCT 392
Db      157  GluMetAlaAlaGlnGlnAsnProLeuGlnProArgGlyArgGlyLeuGlyGln 176
QY      393  CGGGGTACCGCCCGCAGGCTCACCTGTGGCAGCGGGGGCCCGACCGAGCAGACGACAA 452
Db      177  ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY      453  GTGGACATCCCTCTCGGCTCTGCTGGTGGCCACCGACCTATGTGGTGGCTATTGGCAAG 512
Db      194  CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
QY      513  GAGGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCAAAGATAGACGTGCATAGG 572
Db      214  GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY      573  AAGGAGACGCGAGGTGCACCTGAAAAGCCATCAGTGTGCATCTCCACCCCTGAGGGCTGC 632
Db      234  LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
QY      633  TCCTCCGCTGTGTAAGATGATCTTGGAGATTATGCTATAAAGAGGCTAAGGACACCAACG 692
Db      254  SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
```

QY 513 GAGGGGCCCATCCGCAACATCACAAACAGACCCAGTCCAAAGATAGACGTGCATAGG 572
 Db 214 GluGlyAlaThrIleArgAsnIleThrIleGlnThrGlnSerLysIleAspValHisArg 233
 QY 573 AAGGAGAACGAGTGGAGCTGAAAGAACGATCAGTGTGCTACTCCACCCCTGAGGGTGC 632
 Db 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
 QY 633 TCCTCCCTGTAGATCATCTCGAGATATGATATTAAGAGGCTTAAGGACACACAAACG 692
 Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
 QY 693 GCTCAGCAGGTTCCTCCGCAACATCTGCGCCCATATTAATCTTTAGGGCGTCTCATTCGC 752
 Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
 QY 753 AAGGAGGACGGAACCTGAAGAGTAGAGCAGGATACCGAGACAAAATCACCTCC 812
 Db 294 LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
 QY 813 TCGTTCAGACCTTACCTTTACACCTGAGAGGACCATCAGTGTGAGGGGCGCCATC 872
 Db 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
 QY 873 GAGAATTCGTGCGGGCCGAGCAGGAAATATGAAGAAAGTTTCGGGAGCGCTTATGAGAT 932
 Db 334 GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
 QY 933 GATGTGGTGCATGAGC-----TCTCACTGTATCCCTGGCTGAACTGCTGCTGTA 986
 Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeu 373
 QY 987 GGTCTTTTCCGAGCTTCATCCAGCGCAGTCCCGCG-----CCTCCAGCAGCGT 1037
 Db 374 GlyLeuPheProProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
 QY 1038 ACTGGGGTGTCTCCCTAGCTCTCTTATGAGGCTCCGAGCAGGAGTGTGTGAGGTG 1097
 Db 394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
 QY 1098 TTTATCCCGCCGAGCAGTGGCGCATCATCGGCAAGAGGGGAGCAGCATCAACAG 1157
 Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyLysIleLysGln 430
 QY 1158 CTCTCCCGTTTCCGAGCCCTCCATCAAGATTGCACCCGCAACACCTCACTCCAAA 1217
 Db 431 LeuSerArgPheAlaGlyAlaSerIleIleAlaProAlaGluAlaProAspAlaLys 450
 QY 1218 GTTCGTATGTTTATCATCTGACCGCCGAGGCGCCCAATTCAGGCTCAGGGAAGATC 1277
 Db 451 ValArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 470
 QY 1278 TATGGAACACTCAAGAGAGAACTTCTTTGGTCCCAAGGAGGAGTGAAGCTGAGAC 1337
 Db 471 TyrGlyLysIleLysGluLeuAsnPheValSerProLysGluGluValLysLeuGluAla 490
 QY 1338 CACATACGTGTCCACATCAGCAGCTGCGCGTCAATGGCAAGAGTGGAAAAAGGTG 1397
 Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleLysGlyLysThrVal 510
 QY 1398 AACGAGTTCAGAAATTTGACGGCAGCTGAGGTGTAGTGTAGTACCAAGAGCAGACCCCTGAT 1457
 Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
 QY 1458 GAGAACGACAGGTCTCGTGAATAATCCGACATTTCTATCCAGTCAGTGCCTCAA 1517
 Db 531 GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
 QY 1518 CGAAGATCCGACATCTCTGGCCCGCAGTTAAGCAG---CAGCATCAGAGGGA---CAG 1571
 Db 551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570

QY 1572 AGTAACACGCGCCAGGACCGAGGAG 1598
 Db 571 SerGlyProGlnSerArgLys 579
 RESULT 5
 US-09-643-597-176
 ; Sequence 176, Application US/09643597
 ; Patent No. 6426072
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Hosken, Nancy
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Li, Samuel X.
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: McNeill, Patricia D.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.455C11
 ; CURRENT APPLICATION NUMBER: US/09/643,597
 ; CURRENT FILING DATE: 2000-08-21
 ; NUMBER OF SEQ ID NOS: 369
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 176
 ; LENGTH: 579
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-643-597-176
 Alignment Scores:
 Pred. No.: 8,4e-161 Length: 579
 Score: 1873.00 Matches: 380
 Percent Similarity: 84.48% Conservative: 50
 Best local Similarity: 74.66% Mismatches: 63
 Query Match: 53.04% Indels: 16
 DB: 4 Gaps: 9
 US-09-270-437D-7 (1-1946) x US-09-643-597-176 (1-579)
 QY 102 CGGTTAGAGCGCGGAAATTCAAATCCCAATATTCACCCAGCTCCGATGGGAAGTA 161
 Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProHisLeuGlnIleVal 96
 QY 162 CTGGACAGCTCTGCTGCTCAGTATGTACAGTAGAGAACTGTGAGCAAGTGAACACCGAG 221
 Db 97 LeuAspSerLeuLeuValGlnThrGlyValValGluSerCysGluGlnValAsnThrAsp 116
 QY 222 AGTGAGACGCGAGTGTGAATGTCCCTATTCACCTATTCACCGGGAGCAGACCGAGCATC 281
 Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
 QY 282 ATGAAGCTCAATGGCCACCACTGGAGAACCATGCCCTCAAGTCTCTACATCCCGAT 341
 Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaIleThrProAsp 156
 QY 342 GAGCAGATAGCA---CAGGGACCTGAGAATGGCGCGCAGGG-----GCCTTTGGCTCT 392
 Db 157 GluMetAlaAlaGlnGlnAsnProLeuGlnProArgGlyArgGlyLeuGlyGln 176
 QY 393 CGGGGTACGCCCGCAGGGCTCACCTGTGGCAGCGGGGGCCCGCCAGCAGGACGACAA 452
 Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
 QY 453 GTGGACATCCCTTCCTGGCTCTGCTGGTGGCCACCACTGTGGTGGTGGCTATTATTCGCAAG 512
 Db 194 CysAspLeuProLeuArgLeuValProThrGlnPheValGlyAlaIleLysGlyLys 213
 QY 513 GAGGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCCAAGATAGACGTGCATAGG 572

177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
453 GTGGACATCCCTTCCGCTCCTGGTCCGACCCAGCATGATGGTGCATTATTGGCAAG 512
194 CysAspLeuProLeuArgLeuValProThrGlnPheValGlyAlaIleGlyLys 213
513 GAGGGGCCCATCCGCAATCACAACACAGACCCAGTCCCAAGATAGAGTGCATAGG 572
214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
573 AAGGAGAACCGAGTGCAGTGAAGAACCCATCAGTGCACCTCCACCCCTGAGGGCTGC 632
234 LysGluAsnAlaGlyAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
633 TCCTCCCTGTGAAGATGATCTTGGAGATTATGCATTAAGAGGCTAAGACACCAACG 692
254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
693 GCTGACGAGGTTCCCTGAGATCCTGGCCCATTAATCACTTTGTAGGCGCTCATGGC 752
274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
753 AAGGAGAACCGAGTGAAGAACCCATCAGTGCACCTCCACCCCTGAGGGCTGC 812
294 LysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
813 TCCTTGAAGACCTTACCTTTTCAACCTGAGAGGACCATCATCTGTGAAGGGGCGCATC 872
314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
873 GAGAAATGTTGAGGGCCGAGGAGGAATATCAGAAAGTTCGGGAGGCGCTATGAAAT 932
334 GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
933 GATGTGCTGCCATGAGC-----TCTCACCTGATCCCTGGCTGCAACCTGGCTGTGTA 986
354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeu 373
987 GGTCTTTTCCAGCTTCATCCAGGCGAGTCCGCGC-----CCTCCAGCAGCGTT 1037
374 GlyLeuPheProThrSerGlyMetProProThrSerGlyProProSerAlaMet 393
1038 ACTGGGCGCTCCTCCATAGCTCCTTTATCAGGCTCCCGAGGAGATGTCGAGGTG 1097
394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
1098 TTTATCCCGCCAGGCGAGTGGCGCCCATCATCGGCAAGAGGGGAGGAGCATCAACAG 1157
411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlnHisIleLysGln 430
1158 CTCTCCCGTTTCCAGCGCTCATCAAGATTGCAACCCCGCAACACCTGACTCCAA 1217
431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
1218 GTTCGTATGTTATCATCACTGAGCCGAGGCGCCCAATTCAGGCTCAGGAGGAATC 1277
451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
1278 TATGGCAAACTCAAGGAGGAGAACTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGAC 1337
471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
1338 CACATAGTGTGTCAGCATCAGCAGTGGCGGGTCTATTGGCAAAAGTGAACAAACGGTG 1397
491 HisIleArgValProSerPheAlaGlyArgValIleGlyLysGlyLysThrVal 510
1398 AACGAGTTGAGAAATTGAGCGAGCTGAGTGGTAGTACCAAGAGACACGACCCCTGAT 1457
511 AsnGluLeuGlnAsnLeuSerAlaGluValValProArgAspGlnThrProAsp 530
1458 GAGAACGACGAGTTCATCGTGAATAATCATCGACATTTCTATGCCAGTCAGATGGCTCAA 1517
531 GluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550

1518 CGGAAGATCCGAGACATCTCTGGCCCGAGTTAAGCAG---CAGCATCAGAAGGGA---CAG 1571
551 ArgLysIleGlnIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
1572 AGTAACGAGCCCGCAGCAGCAGGAGAG 1598
571 SerGlyProGlnSerArgArgLys 579

RESULT 4

US-09-606-421B-348
; Sequence 348, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A. W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-606-421B-348

Alignment Scores:
Pred. No.: 5,54e-161 Length: 579
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best Local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 4 Gaps: 9

US-09-270-437D-7 (1-1946) x US-09-606-421B-348 (1-579)

102 CGGGTTAGAGCGCGGAAATTCAAATCCGAAATATTCCACCCAGCTCCGATCGGAAGTA 161
77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluVal 96
162 CTGGACAGCTGCTGCTCAGTATGTA CAGTAGAAGTGTGAGCAAGTGAACACCGAG 221
97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
222 AGTGACAGCGCATGTGTGATCTACCTATTCCACCGGAGCAGACACGAGCAGCATC 281
117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaLeu 136
282 ATGAACCTCAATGGCCACCGAGTGGAGAACCATGCCCTGAAGTCTCTACATCCCCGAT 341
137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
342 GAGCAGATAGCA---CAGGGACCTGAGAAATGGGCGCCGAGGG-----GGCTTTGGCTCT 392
157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGln 176
393 CGGGGTGACCCCGCCAGGCGCTCACCCTGTGGCAGCGGGGGCCCGCCAGCAGCAGCAA 452
177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
453 GTGGACATCCCTTCCGCTCCTGGTGGTGCCACCCAGTATGTGGTGCCATTATTGGCAAG 512
194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213

Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
 QY 342 GAGCAGATAGCA--CAGGACCTGAGAAATGGGCGCGAGGG-----GGCTTTGGCTCT 392
 Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGln 176
 QY 393 CGGGGTGAGCCGCCAGGGCTACCTGTGTCAGCGGGGGCCCGACCCAGCAGCAGCAA 452
 Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
 QY 453 GTGACATCCCTCCCTCGCTCTGGTGGCCAGCCAGTATGTGGTGCATATTGGCAAG 512
 Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
 QY 513 GAGGGGCGCCATCCGCAACATCACAAACAGACCCAGTCCCAAGATAGAGTGCATAGG 572
 Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
 QY 573 AAGGAGAACCGAGTGCAGCTGAAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGC 632
 Db 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
 QY 633 TCCTCCGCTTTAAGATGATCTCGAGATTATGCATTAAGAGGCTTAAGGACACCAAAACG 692
 Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
 QY 693 GCTGACGAGGTTCCCTCAAGATCTCGCCCATATAACTTTGTAGGGCGTCTCATTCGC 752
 Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
 QY 753 AAGGAAGCAGGAACCTGAAGAGGTAGAGAAGATACCGAGACAAATAATCATCTCC 812
 Db 294 LysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
 QY 813 TCGTTGACAGCTTACCTTTACACCTGAGAGCCATCAGTGTGAAGGGGCCATC 872
 Db 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
 QY 873 GAGAAATCTGTCAGGGCGCAGCAGCAAAATATGAAGAAGTTCCGGAGGCTTATGAGAA 932
 Db 334 GluThrCysAlaLysAlaGluGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
 QY 933 GATGTGGTGCATGAGC-----TCTCACTGTATCCCTGGCTGAACTGCTGCTGTA 986
 Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeu 373
 QY 987 GGTCTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCTCCAGCAGCGTT 1037
 Db 374 GlyLeuPheProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
 QY 1038 ACTGGGGCTGCTCCCTATAGCTCTCTTTATGCAAGGCTCCCGAGCAGAGATGTCGAGGTG 1097
 Db 394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
 QY 1098 TTTATCCCGCCAGCGAGTGGCGCCATCATCGGCAAGAGGGGAGCAGACATCAACAG 1157
 Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGlnHisLysGln 430
 QY 1158 CTCTCCCGTTTGGCAGCGCTTCATCAGATTGACACCCGCAACACCTGACTCCCAA 1217
 Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
 QY 1218 GTTCTATGTTATCATCTAGCGCCGCGCAGAGGCGCCAAATTCAGAGCTCAGGAGAAATC 1277
 Db 451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
 QY 1278 TATGGCAACTCAAGAGAGAACTTCTTTGTTGTCCTCCAGAGAGAGTGAAGCTGGAGCC 1337
 Db 471 TyrGlyLysIleLysGluLysGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
 QY 1338 CACATACGTGTCCAGCATCAGCAGTGGCGGCTCATTTGGCAAGAGTGGAAACCGGTG 1397
 Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510

QY 1398 AACAGATTGAGAAATTTGACCGCAGCTGAGTGGTAGTACCAAGAGACAGACCCCTGAT 1457
 Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
 QY 1458 GAGAACACACAGGTTCATCGTGAATAATCATCGACATTTCTATGTCAGTCCAGATGCTCAA 1517
 Db 531 GluAsnAspGlnValValValIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
 QY 1518 CGAAGATCCGAGACATCTCGCCCGAGGTTAAGCAG---CAGCATCAGAAGGGA---CAG 1571
 Db 551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
 QY 1572 AGTAACACGCGCCAGGACGAGGAGAG 1598
 Db 571 SerGlyProProGlnSerArgArgLys 579

RESULT 3

US-09-542-615A-348
 ; Sequence 348, Application US/09542615A
 ; Patent No. 6518256
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Hosken, Nancy A.
 ; APPLICANT: Fanger, Gary R.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121-455C8
 ; CURRENT APPLICATION NUMBER: US/09/542,615A
 ; CURRENT FILING DATE: 2000-04-14
 ; NUMBER OF SEQ ID NOS: 350
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 348
 ; LENGTH: 579
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-542-615A-348

Alignment Scores:
 Pred. No.: 5,54e-161 Length: 579
 Score: 1875.00 Matches: 380
 Percent Similarity: 84.68% Conservative: 51
 Best Local Similarity: 74.66% Mismatches: 62
 Query Match: 53.10% Indels: 16
 DB: 4 Gaps: 9

US-09-270-437D-7 (1-1946) x US-09-542-615A-348 (1-579)

QY 102 CGGGTTAGGAGCGGAAATTCAAATCCGAAATATCCACCCAGCTCCGATGGGAAGTA 161
 Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluVal 96
 QY 162 CTGGACAGCTCTGGCTCAGTATGTCAGTACAGACTGTGAGCACTGTGACCACTGACACCGAG 221
 Db 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
 QY 222 AGTGAGAGCGGAGTGGTGAATCTCACTTATCCACCGGAGCAGACCCAGGCAAGCCATC 281
 Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
 QY 282 ATGAAGCTGAATGGCCACAGTTGGAGAACCAATGCCCTGAAGGTCTCTACATCCCGAT 341
 Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
 QY 342 GAGCAGATAGCA--CAGGACCTGAGATGGCGCGCGAGGG-----GCTTTGGCTCT 392
 Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGln 176
 QY 393 CGGGGTGAGCCCGCGCAGGGCTCACCTGTGGCAGCGGGGGCGCCAGCAGCAGCAA 452

QY 228 ACAGCATGTGTAATGTCACCTATTCCAAACCGGAGCAGACCGAGCAAGCCATCATGAAG 287
Db ThrAlaValValAsnValThrTyrSerAsnArgGluGlnThrArgGlnAlaIleMetLys 138
QY 288 CTGATGGCCACCACTGTGGAGAACCATGCCCTGAAGTCTCTCATATCCCGATGAGCAG 347
Db LeuAsnGlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyrIleProAspGluGln 158
QY 348 ATAGCACAGGACCTGAGAAATGGCGCGCGAGGGGGCTTTGGCTCTCCGGGGTCACCCCGC 407
Db IleThrGlnGlyProGluAsnGlyArgGlyGlyPheGlySerArgGlyGlnProArg 178
QY 408 CAGGGCTCACCTGTGGCAGCGGGGGCCCGCCAGCAGCAGCAAGTGGACATCCCCCTT 467
Db GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnProValAspIleProLeu 198
QY 468 CGGCTCTGTGGCCACCCAGTATGTGGTGCCATTATTGGCAAGGAGGGGGCCACCATC 527
Db ArgLeuLeuValProThrGlnTyrValGlyAlaIleGlyLysGluGlyAlaThrIle 218
QY 528 CGCAACATCAAAACAGACCCAGTCCCAAGATAGACGTGCATAGAGAGAGAAACGACGGT 587
Db ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGly 238
QY 588 GCAGCTGAAAAGCCATCAGTGTGCATCCACCCCTGAGGGCTCTCTCCCGTTGTAAG 647
Db AlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCysSerSerAlaCysLys 258
QY 648 ATGATCTTGGAGATTATGCATAAAGAGCTAAAGACACCAAAACGGCTGACGAGTTCCC 707
Db MetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThrAlaAspGluValPro 278
QY 708 CTGAAGATCTGGCCCCATAATACTTTGTAGGGCGTCTCATTTGGCAAGAGAGACGGAAC 767
Db LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn 298
QY 768 CTGAGAGGTAGACAGATACCGACACAAATACCATCTCTCTGTTGCAAGACCTT 827
Db LeuLysLysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeu 318
QY 828 ACCCTTTACAACTGAGAGACCATCATCTGTAAGGGGGCCATCGAGAAATTTGTCAGG 887
Db ThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIleGluAsnCysArg 338
QY 888 GCGCAGCAGGAAAATAATCAAGAAAGTTCCGAGGCGCTTATGAGATGATGTGGCTGCATG 947
Db AlaGluGlnGluIleMetLysLysValArgGluAlaTyrGluAsnAspValAlaAlaMet 358
QY 948 AGC-----TCTACCTGTATCCCTGGCGTGAACCTGCTGCTGTAGTCTTTTCCAGCT 1001
Db SerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaValAlaGlyLeuPheProAla 378
QY 1002 TCATCCAGCGCAGTCCCGCGCTCCAGCAGCTTACTGGGGCTGTCTCCCTATAGCTCC 1061
Db SerSerSerAlaValProProProSerSerValThrGlyAlaAlaProTyrSerSer 398
QY 1062 TTATGACGGCTCCGAGCAGGAGATGTGCAGGTGTTATCCCGCGCCACGAGCAGTGGGC 1121
Db PheMetGlnAlaProGluGlnGluMetValGlnValPheIleProAlaGlnAlaValGly 418
QY 1122 GCATCATCGCAGAGGGGCAGCAGATCAACAGCTCTCCCGGTTTGGCAGGCGCTCC 1181
Db AlaIleIleGlyLysLysGlyGlnHisIleLysGlnLeuSerArgPheAlaSerAlaSer 438
QY 1182 ATCAAGATTGCACACCGCAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGGA 1241
Db IleLysIleAlaProProGluThrProAspSerLysValArgMetValAlaIleThrGly 458
QY 1242 CCGCCAGAGGCCCAATTCAGGGCTCAGGGAAGAATCTATGCAAACTCAAGGAGGAGAAC 1301
Db ProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysLeuLysGluGluAsn 478

QY 1302 TTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCA 1361
Db PhePheGlyProLysGluGluValLysLeuGluThrHisIleArgValProAlaSerAla 498
QY 1362 GCTGGCCGGGTCTATGGCAAGGTGGAAAAACGGTGAACGAGTTGCAGAAATTTCACGGCA 1421
Db AlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThrAla 518
QY 1422 GCTGAGGTGTAGTACCACAGAGACACACCCCTGTAGAGAACACACAGGTTCATCGTCAAA 1481
Db AlaGluValValValProArgAspGlnThrProAspGluAsnAspGlnValIleValLys 538
QY 1482 ATCATCGGACATTTCTATGTCAGTGCAGATCGCTCAACGGAAGATCCGAGACATCTCGCC 1541
Db IleIleGlyHisPheTyrAlaSerGlnMetAlaGlnArgLysIleArgAspIleLeuAla 558
QY 1542 CAGGTTAAAGCAGCAGCATCAGAGGAGCAGAGTAACAGGCCCGCAGGACCGAGGAAG 1598
Db GlnValLysGlnGlnHisGlnLysGlyGlnSerAsnLeuAlaGlnAlaArgLys 577

RESULT 2
US-09-643-597-348
; Sequence 348, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-643-597-348

Alignment Scores:
Pred. No.: 5,54e-161 Length: 579
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best Local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 4 Gaps: 9

US-09-270-437D-7 (1-1946) x US-09-643-597-348 (1-579)

QY 102 CGGGTTAGGAGCCGGAAATTCGAATCCGAATATTCACCCCGCTCCGATGGAGTA 161
Db ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluVal 96
QY 162 CTGCAGACGCTGTGCTCAGTATGTGTACAGTAGAAGTGTGAGCAAGTGAACACCGAG 221
Db LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGACAGCGCAGTGTGTGAATGTCACTATTCCACCGGAGCAGACCGAGCAAGCCATC 281
Db SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATCAAGCTGAATGCGCCACCAAGTTGGAGAACCATGCCCTGAAGTCTCTCATCCCGAT 341

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 23, 2004, 10:56:36 ; Search time 17.4913 Seconds
(without alignments)
11487.328 Million cell updates/sec

Title: US-09-270-437D-7
Perfect score: 3531
Sequence: 1 gctgtagggagggtggg.....atttccttcagggttttaaaa 1946

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 778828

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US0270437/runat_23072004_095740_27466/app_query.fasta_1.5582
-DB=Issued Patents AA -QFMT=fastan -SURFIX=ral -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0270437@cgn 1.1 59 @runat_23072004_095740_27466 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:
1: /cgn2_6/prodata/2/iaa/5A COMB.pcp:*
2: /cgn2_6/prodata/2/iaa/5B COMB.pcp:*
3: /cgn2_6/prodata/2/iaa/6A COMB.pcp:*
4: /cgn2_6/prodata/2/iaa/6B COMB.pcp:*
5: /cgn2_6/prodata/2/iaa/ECTUS COMB.pcp:*
6: /cgn2_6/prodata/2/iaa/backfiles1.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2491	70.5	577	3	US-09-261-855-2
2	1875	53.1	579	4	US-09-643-597-348
3	1875	53.1	579	4	US-09-542-615A-348
4	1875	53.1	579	4	US-09-606-421B-348
5	1873	53.0	579	4	US-09-643-597-176
6	1873	53.0	579	4	US-09-480-884A-176
7	1873	53.0	579	4	US-09-542-615A-176
8	1873	53.0	579	4	US-09-606-421B-176
9	248.5	7.0	643	1	US-08-021-608D-10
10	248.5	7.0	643	1	US-08-726-160-10
11	248.5	7.0	643	5	PCT-US94-01782-10
12	245	6.9	49	3	US-09-261-855-18

13	245	6.9	644	1	US-08-021-608D-2	Sequence 2, Appli
14	245	6.9	644	1	US-08-726-160-2	Sequence 2, Appli
15	245	6.9	644	5	PCT-US94-01782-2	Sequence 2, Appli
16	242	6.9	590	1	US-08-021-608D-8	Sequence 8, Appli
17	242	6.9	590	1	US-08-726-160-8	Sequence 8, Appli
18	242	6.9	590	5	PCT-US94-01782-8	Sequence 8, Appli
19	236	6.7	48	3	US-09-261-855-20	Sequence 20, Appli
20	229	6.5	47	3	US-09-261-855-17	Sequence 17, Appli
21	228	6.5	47	3	US-09-261-855-19	Sequence 19, Appli
22	222	6.3	49	3	US-09-261-855-22	Sequence 22, Appli
23	221.5	6.3	530	1	US-08-187-793-4	Sequence 4, Appli
24	219	6.2	48	3	US-09-261-855-24	Sequence 24, Appli
25	214	6.1	47	3	US-09-261-855-21	Sequence 21, Appli
26	191	5.4	47	3	US-09-261-855-23	Sequence 23, Appli
27	174.5	4.9	1418	4	US-09-252-991A-32367	Sequence 32367, A
28	174	4.9	343	1	US-08-187-793-2	Sequence 2, Appli
29	169.5	4.8	577	4	US-09-252-991A-19178	Sequence 19178, A
30	169	4.8	720	4	US-09-252-991A-21881	Sequence 21881, A
31	165	4.7	657	4	US-09-252-991A-28001	Sequence 28001, A
32	162.5	4.6	489	4	US-09-252-991A-16940	Sequence 16940, A
33	162.5	4.6	697	4	US-09-252-991A-24009	Sequence 24009, A
34	160	4.5	774	4	US-09-252-991A-16789	Sequence 16789, A
35	159	4.5	1008	4	US-09-252-991A-29419	Sequence 29419, A
36	157	4.4	782	4	US-09-252-991A-30464	Sequence 30464, A
37	154.5	4.4	243	1	US-08-021-608D-6	Sequence 6, Appli
38	154.5	4.4	243	1	US-08-726-160-6	Sequence 6, Appli
39	154.5	4.4	243	5	PCT-US94-01782-6	Sequence 6, Appli
40	154.5	4.4	904	4	US-09-976-594-615	Sequence 615, App
41	153	4.3	1093	4	US-09-252-991A-21827	Sequence 21827, A
42	151	4.3	1411	4	US-09-252-991A-28408	Sequence 28408, A
43	150.5	4.3	957	4	US-09-252-991A-20408	Sequence 20408, A
44	149	4.2	1706	4	US-09-252-991A-31760	Sequence 31760, A
45	148.5	4.2	563	4	US-09-252-991A-28458	Sequence 28458, A

ALIGNMENTS

RESULT 1

US-09-261-855-2
; Sequence 2, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-2

Alignment Scores:
Pred. No.: 1.01e-216 Length: 577
Score: 2491.00 Matches: 493
Percent Similarity: 99.00% Conservative: 1
Best Local Similarity: 98.80% Mismatches: 3
Query Match: 70.55% Indels: 2
DB: 3 Gaps: 1

US-09-270-437D-7 (1-1946) x US-09-261-855-2 (1-577)

Qy	108	AGGAGCGGAAATTCGAAATATTCACCCAGCTCCGATGGGAGTACTGGAC	167
Db	79	ArgSerArgLysIleGlnIleArgAsnIleProGlnLeuArgTrpGluValLeuAsp	98
Qy	168	AGCTGCTGCTCAGTATGTCAGTAGAGAACTGTGAGCAAGTGAACACCGAGTGG	227
Db	99	SerLeuLeuAlaGlnTyrGlyThrValGluAsnCysGluGlnValAsnThrGluSerGlu	118

This Page Blank (uspto)

Job time : 112.41 secs

```
|||||
234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
QY TCCTCCGCTGCTTAAGATGATCTTCGAGATTATGATAAAGAGGCTTAAGACACCAAAAACG 692
Db :|||
254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
QY GCTACAGAGGTTCCCTCAAGATCCTGGCCCATATAACTTTGTAGGGCGTCTCATGGC 752
Db :|||
274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY :|||
753 AAGAAGACGAGCACTGAAGAGTAGACAGACAGATACCGAGACGACAAATCACCATCTCC 812
Db :|||
294 LysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
QY :|||
813 TC GTTGAAGACCTTACCTTTACACCTGAGAGGACCATCACTGTGAGGGGGCCATC 872
Db :|||
314 ProLeuGlnGluLeuThrLeuTyraAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY :|||
873 GAGAAATTGTTGACGGCGGACGAGCAATAATGAAGAAAGTTGCGGAGCGCTATGAGAAAT 932
Db :|||
334 GluThrCysAlaLysAlaGluGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
QY :|||
933 GATGTGCTGCATGAGC-----TCTCACCTGATCCCTGGCTGAACCTGGCTGCTGTA 986
Db :|||
354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeu 373
QY :|||
987 GGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCTCCAGCAGCGCTT 1037
Db :|||
374 GlyLeuPheProProThrSerGlyMetProProThrProThrSerGlyProProSerAlaMet 393
QY :|||
1038 ACTGGGCTGCTCCCTATAGCTCCTTTATGAGGCTCCCGAGCAGAGATGGTGAGGTG 1097
Db :|||
394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisGln 410
QY :|||
1098 TTTATCCCGCGCAGGCGAGTGGGCGCATCATCCGCAAGAGAGGGCGCAGCATCAACAG 1157
Db :|||
411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGln 430
QY :|||
1158 CTCTCCCGGTTTGGCAGGCGCTCCATCAAGATTGCACCAACCCGCAACACCTGACTCCAAA 1217
Db :|||
431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY :|||
1218 GTTCGTATGTTATCATCATCGACCGCCAGAGGCCCAATTCAGGCTCAGGGAAGAATC 1277
Db :|||
451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY :|||
1278 TATGGCAACTCAAGGAGAGAACTTCTTTGGTCCCAAGGAGGAGTGAAGCTGGAGACC 1337
Db :|||
471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY :|||
1338 CACATACGTGTGCCAGCATCAGACGTGCGCGGTTCATTGGCAAGAGTGGAACACGGTG 1397
Db :|||
491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyGlyLysThrVal 510
QY :|||
1398 AACGAGTTGCAGAAATTCACGCGACGCTGAGTGGTAGTACCAAGAGACAGACCCCTGAT 1457
Db :|||
511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
QY :|||
1458 GAGAACGACCGAGGTTCATCGTGAATATCATCGACATTTCTATGCCAGTCCAGTGGCTCAA 1517
Db :|||
531 GluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
QY :|||
1518 CGGAAGATCCGAGACATCCTGGCCCGAGGTTAAGCAG---CAGCATCAGAAGGGA---CAG 1571
Db :|||
551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnLysAlaLeuGln 570
QY :|||
1572 AGTAACCGAGCGCGAGGAGGAG 1598
Db :|||
571 SerGlyProProGlnSerArgArgLys 579
```

```
QY 933 GATGTGGCTGCCATGAGC-----TCTCACTGATCCCTGGCCTGAACCTGGCTGCTGTA 986
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 361 AspileAlaSerMetLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeu 380
QY 987 GGTCTTTTCCAGCTTCATCCAGCGAGTCCTCCGCG-----CCTCCAGCAGGTT 1037
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 381 GlyLeuPheProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 400
QY 1038 ACTGGGGCTCTCCCTATAGCTCTCTTATCAGGCTCCCGAGCAGGAGATGGTGCAGGTG 1097
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 401 Thr-----ProProlyrProGlnPheGluGluSer---GluThrGluThrValHisLeu 417
QY 1098 TTTATCCCGCCCGCAGCGAGTGGCGCATCATCGCAAGAGGGCGAGCAGCATCAACAG 1157
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 418 PheileProAlaLeuSerValGlyAlaIleGlyGlnGlyGlnHisIleLysGln 437
QY 1158 CTCTCCGGTTTCCAGCGCTTCATCAGATTCGACCCAGGAGGAGGAGGAGGAGGAGGAG 1217
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 438 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 457
QY 1218 GTTCGTATGTTATCATCTGACCGCCAGCGAGGCCCAATTCAAGGCTCAGGGAAGATC 1277
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 458 ValArgMetValIleleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 477
QY 1278 TATGGCAAACTCAAGGAGGAGAACTTTCTTTGGTCCCAAGGAGGAGGAGGAGGAGGAG 1337
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 478 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 497
QY 1338 CACATAGGTGTGCAGCATCAGCAGCTGGCGGGTCAATTGGCAAGGTGGAAGAGGAGTG 1397
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 498 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 517
QY 1398 AACAGAGTGCAGATTTGACCGCAGCTGAGGTGTAGTACCAAGAGCAGCAGCCCTGAT 1457
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 518 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgaspGlnThrProasp 537
QY 1458 GAGAACGACAGGTCTATCGTGAATAATCATCGACATTTCTATGCGCAGTCAAGTGGCTCAA 1517
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 538 GluAsnAspGlnValValValLysIleThrGlyHisPheThrAlaCysGlnValAlaGln 557
QY 1518 CGNAGATCCGAGACATCTCGGCCCGGTTAAGCAG---CAGCATCAGAGGGA---CAG 1571
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 558 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 577
QY 1572 AGTAACCAAGCCAGGCGCAGCAGGAGAG 1598
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 578 SerGlyProProGlnSerArgArgLys 586

RESULT 15
ID AAB11328
XX
AC AAB11328;
XX
DE 21-FEB-2001 (first entry)
XX
KW Human lung cancer-associated protein L523S.
XX
KW Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;
XX
KW vaccine; detection.
XX
OS Homo sapiens.
XX
PN WO200061612-A2.
XX
PD 19-OCT-2000.
XX
PF 03-APR-2000; 2000WO-US008896.
XX
PR 02-APR-1999; 99US-00285479.
PR 17-DEC-1999; 99US-00466396.
PR 30-DEC-1999; 99US-00476496.
```

```
PR 10-JAN-2000; 2000US-00480884.
PR 22-FEB-2000; 2000US-00510376.
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Fan L;
XX
DR WPI; 2000-628399/60.
DR N-PSDB; AAC65900.
XX
PT Isolated polypeptide comprising an immunogenic portion of a lung tumor
PT protein is used for detecting and monitoring progression of lung cancer
PT in a patient.
XX
XX Claim 3; Page 186-188; 261pp; English.
XX
CC This invention describes a novel isolated polypeptide (I) which
CC comprising an immunogenic portion of a lung tumor protein or variant (P2)
CC which have cytostatic activity. The polypeptides and polynucleotides are
CC used in compositions and vaccines to inhibit the development of cancer,
CC especially lung cancer, in a patient. Methods described in the invention
CC can be used to monitor the progression of a cancer by carrying out the
CC detection at subsequent time points and comparing the results from the
CC different time points. CD4+ and/or CD8+ T-Cells isolated from a patient
CC are treated with P2, polynucleotides encoding P2 or antigen presenting
CC cells expressing P2, and then administered to the patient to inhibit
CC development of cancer
XX
SQ Sequence 579 AA;

Alignment Scores:
Pred. No.: 9,35e-160 Length: 579
Score: 1873.00 Matches: 380
Percent Similarity: 84.48% Conservative: 50
Best Local Similarity: 74.66% Mismatches: 63
Query Match: 53.04% Indels: 16
DB: Gaps: 9

US-09-270-437D-7 (1-1946) x AAB11328 (1-579)
QY 102 CGGGTTAGGAGCGCGAAATTCAAATCCGAAATATTCACCCCGAGCTCCGATGGGAAGTA 161
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnThrGluVal 96
QY 162 CTGGACAGCTCTGCTGCTCAGTATGTAGTACAGAACTGTGAGCAAGTGAACACCGAG 221
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGACAGCGCAGTGTGTGATGTCACTATTCACCGGAGCAGCAGCAGGAGCCATC 281
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATGAAGTGAATGGCCACCGAGTGGAGAACCATGCCCTGAAGTCTCTACATCCCGCAT 341
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY 342 GAGCAGATAGCA---CAGGACCTGAGAAATGGGCGCCGAGGG-----GGCTTTGGCTCT 392
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 157 GluMetAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGln 176
QY 393 CGGGGTTCAGCCCGCAGGGCTCACCTGTGGCAGCGGGGGCCCGCAGCAGCAGCAA 452
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 493
QY 453 GTGGACATCCCTTCGGCTCTCGGTGCCACCCAGTATGTGGTGCCATTATTGGCAAG 512
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleLeuGlyLys 213
QY 513 GAGGGGGCCACCATCCCGCAACATCAACAAACAGACCCAGTCCCAAGATAGACGTGATAG 572
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY 573 AAGGAGAACGCGAGGTGCAGCTGAAAGGCCATCAGTGTGCATCCACCCCTGAGGGCTGC 632
```


DR N-PSDB; ABL49283.
 XX Polynucleotides encoding lung tumor polypeptides, useful for treating
 PT lung cancer or stimulating an immune response.
 PS
 XX Claim 2; Page 354-355; 374pp; English.
 XX The present invention describes human lung tumour proteins. Human lung
 CC tumour proteins and polynucleotides have cytostatic and immunostimulant
 CC activities, and can be used in vaccine production. Compositions
 CC comprising the lung tumour proteins, polynucleotides, antibodies, fusion
 CC proteins, T cell populations, or antigen presenting cells that express
 CC the lung tumour proteins are useful for treating lung cancer or
 CC stimulating an immune response. ABL48959 to ABL49300 and ABL49496 to
 CC ABL50700 represent sequences used in the exemplification of the present
 CC invention
 XX
 SQ Sequence 586 AA;

Alignment Scores:
 Pred. No.: 6.2e-160 Length: 586
 Score: 1875.00 Matches: 380
 Percent Similarity: 84.68% Conservative: 51
 Best Local Similarity: 74.66% Mismatches: 62
 Query Match: 53.10% Indels: 16
 DB: 5 Gaps: 9

US-09-270-437D-7 (1-1946) x ABL75048 (1-586)

102 CGGTTAGGAGCGGAAATTCMAATCCGAATATTCACCCAGCTCCGATGGAGTA 161
 Db 84 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluVal 103
 162 CTGGACAGCTGCTGGCTCAGTATGTACAGTAGAGAACTGTGAGCAAGTGAACACCGAG 221
 Db 104 LeuAspSerLeuLeuValGlnTrpGlyValValGluSerCysGluGlnValAsnThrAsp 123
 222 AGTCAGAGCGGAGTGGTGAATGTACCTATTCCAAACCGGAGCAGACCGCAAGCCATC 281
 Db 124 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 143
 282 ATGAAGCTGAATGGCCACAGTTGGAGAACCCATGCCCTGAAGTCTCTACATCCCGAT 341
 Db 144 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 163
 342 GAGCAGATAGCA---CAGGGACCTGAGAATGGCGCCGAGGG-----GGCTTTGGCTCT 392
 Db 164 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGln 183
 393 CGGGTTCAGCCCGCAGGGCTCACCTGTGGCAGCGGGGGCCCCCAGCCCAAGCAGCAAA 452
 Db 184 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 200
 453 GTGGACATCCCTTCGGCTCTGGTGGCCACCCAGTATGTGGTGCCATTTATTCGCRAG 512
 Db 201 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 220
 513 GAGGGGGCCACCATCCGCAACATCACAAACAGACCCAGTCCCAAGATAGACGTGATAGG 572
 Db 221 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 240
 573 AAGGAGAACGCGAGGTGCAGCTGAAAAGCCATCAGTGTGCATCCACCCCTCAGGGCTGC 632
 Db 241 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 260
 633 TCCTCGCTGTGAAGATGATCTGGAGATTATGAGCTAAGAGCTAAGGACACCCAAACG 692
 Db 261 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 280
 693 GCTGACGAGGTTCCTCGAAGTCTCGGCCCATATACTTTGTAGGGGCTCTCAATTGCG 752
 Db 281 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 300

753 AAGGAAGGACGGAACCTGAAGAGTAGACAGATACCGAGACAAATAATCACCATCTCC 812
 Db 301 LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer 320
 813 TCGTTGCAAGACCTTACCTTTTCAACCCCTCAGAGGACCATCCTGTGAAGGGGCCATC 872
 Db 321 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 340
 873 GAGAAATGTTTCAGGGCGGACGAGCAATAATATGAAGAAGTTCCGGAGCCCTATCAGAT 932
 Db 341 GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyrGluAsn 360
 933 GATGTGGTGCACAGC-----TCTCACCTGATCCCTGGCTGAACCTGCTCTCTGTA 986
 Db 361 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeu 380
 987 GGTCTTTTCCAGCTTTCATCCAGCGCAGTCCCGCGG-----CCTCCAGCAGCGTT 1037
 Db 381 GlyLeuPheProProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 400
 1038 ACTGGGGTGTCTCCTATATAGCTCTTTATGAGGCTCCGAGCAGGAGATGTCGAGGTG 1097
 Db 401 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu 417
 1098 TTTATCCCGCGCCAGGAGTGGCGCCATCATCGCAAGAGAGGGCAGCACATCAAAACAG 1157
 Db 418 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyHisIleLysGln 437
 1158 CTCTCCCGGTGTGGCAGCGCTTCCATCAAGATTGCACCCAGCAACCTGACTCCAAA 1217
 Db 438 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 457
 1218 GTTGTATGTTATCATCTGACCGCCGAGAGGCCCAATTCAGGCTCAGGGAAGATC 1277
 Db 458 ValArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 477
 1278 TATGCCAACTCAAGGAGGAACCTTCTTGTGCCAAGGAGAGTGAAGCTGGAGACC 1337
 Db 478 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 497
 1338 CACATACGTGTGCCAGCATCAGCAGCTGGCGGGTCAATGGCAAAAGGTGAAAAACGGTG 1397
 Db 498 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 517
 1398 AACGAGTTCGAAATTCAGCGCAGCTGAGGTGTGTACCAAGAGCAGACCCCTGAT 1457
 Db 518 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 537
 1458 GAGAACGACCGAGTTCATCGTGAATATCATCGGACATTTCTATGCCAGTCAGATGGCTCAA 1517
 Db 538 GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 557
 1518 CGAAAGATCCGAGACATCTCGCCCGCAGGTTAAGCAG---CAGCATCAGAAGGGA---CAG 1571
 Db 558 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 577
 1572 AGTAACGAGCGCCAGGACCGAGGAAG 1598
 Db 578 SerGlyProProGlnSerArgArgLys 586

RESULT 13

ABP61968
 ID ABP61968 standard; protein; 586 AA.

XX
 AC ABP61968;

DT 07-OCT-2002 (first entry)

XX Human lung cancer associated protein sequence SEQ ID NO:427.

XX Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.

XX
 OS Homo sapiens.

Best Local Similarity: 74.66%			Mismatches: 62		
Query Match: 53.10%			Indels: 16		
DB: 7			Gaps: 9		
US-09-270-437D-7 (1-1946) x ADA28438 (1-579)					
QY	102	CGGGTTAGAGCGGGAATTCGAATATTCACCCAGCTCCGATGGGAAGTA	161		
Db	77	ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluVal	96		
QY	162	CTGGACAGCGCTGCTGGCTCAGTATGTACAGTAGAAGCTGTGACAGAGTGAACACCGAG	221		
Db	97	LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp	116		
QY	222	AGTGAGAGCGCAGTGTGAATGTACCTATTCACCCGAGGAGCAGCAGGCAAGCCATC	281		
Db	117	SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu	136		
QY	282	ATGAGCTGAATGGCCACACAGTTGGAGAACATGCCCTGAGGTCTCTCATCCCCGAT	341		
Db	137	AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp	156		
QY	342	GAGCAGATAGCA---CAGGACCTGAGAAATGGGCCCGAGGG-----GGCTTTGGCTCT	392		
Db	157	GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGln	176		
QY	393	CGGGGTGAGCCCGCCAGGCTCACCTGTGGCAGCGGGGCCCCAGCCAGCAGCAAA	452		
Db	177	ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro	193		
QY	453	GTGACATCCCTTCCGCTCGCTGCTGCCACCCAGTATGTGGTGCCATATTGGCAAG	512		
Db	194	CysAspLeuProLeuArgLeuValProThrGlnPheValGlyAlaIleIleGlyLys	213		
QY	513	GAGGGGCCACCATCCGCAACATCACAAACAGACCCAGTCCAGATAGACGTGCATAGG	572		
Db	214	GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg	233		
QY	573	AAGGAGACCGAGTGCGAGTGAAGAAAGCCATCATGTGTCACCTCCACCCCTGAGGGTGC	632		
Db	234	LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr	253		
QY	633	TCCTCCGCTTGAAGATGATCTGGAGTATTCGATTAAGAGCGTAAAGCACCAAAACG	692		
Db	254	SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe	273		
QY	693	GCTGACGAGTTCCTCCCTGAGATCTGCGCCCAATAAATTTGTAGGGCTCTCATGGC	752		
Db	274	ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly	293		
QY	753	AAGGAGGAGCGGAACCTGAGAGAGGTAGACAGATACCGAGACAAATAACCATCTCC	812		
Db	294	LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer	313		
QY	813	TCGTGTCAAGACTTACCTTTACACCTGAGAGGACCATCACTGTGAAGGGGCCATC	872		
Db	314	ProLeuGlnGluThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal	333		
QY	873	GAGAAATGTTGCGAGGCGGAGGAGAAATAATGAAGAAAGTTCGGAGGCGCTATGAAAT	932		
Db	334	GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyrGluAsn	353		
QY	933	GATGTGCTGCATGAGC-----TCTCAGCTGATCCCTGGCTGAACTGGCTGTGTA	986		
Db	354	AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeu	373		
QY	987	GGTCTTTTCCAGCTTTCACGCGCAGTGTCGCGCG-----CCTCCCGAGCGGTT	1037		
Db	374	GlyLeuPheProProThrSerGlyMetProProThrSerGlyProProSerAlaMet	393		
QY	1038	ACTGGGCTGCTCCCTATAGCTCTTTATGAGGCTCCCGAGGAGAGATGGTGCAGGTG	1097		
Db	394	Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu	410		
QY	1098	TTTATCCCGCCCGAGCAGTGGCGCCATCATCGCAAGAGGGGAGGACATCAACAG	1157		
Db	411	PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlnHisIleLysGln	430		
QY	1158	CTCTCCCGTTTCCGAGCGCTCCATCAGATTGCACCCGAAACCAACCTGACTCCAAA	1217		
Db	431	LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys	450		
QY	1218	GTTCTGATGTTATCATCTGACCGCAGGAGCCCAATTCAAGGCTCAGGGAAGATC	1277		
Db	451	ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle	470		
QY	1278	TATGGCAACTCAAGGAGGAGAACTTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGAC	1337		
Db	471	TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla	490		
QY	1338	CACATACGTGTGCGACATCAGCAGCTGGCGGGTCTATTGGCAAGAGTGAAGAGCGTG	1397		
Db	491	HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal	510		
QY	1398	AACGAGTTGCAAAATTTGACCGCAGCTGAGGTGGTAGTACCAAGACAGACCCCTGAT	1457		
Db	511	AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgaspGlnThrProasp	530		
QY	1458	GAGACGACGAGCTCATCGTGAATAATCATCGGCATTTCTATGCCAGTGCATGGCTCAA	1517		
Db	531	GluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln	550		
QY	1518	CGGAAGATCCGACACATCTGCGCCAGGTAAAGCAG---CAGCATCAGAGGGA---CAG	1571		
Db	551	ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnLysAlaLeuGln	570		
QY	1572	AGTAACCGAGCCCGAGCGAGGAGGAAG	1598		
Db	571	SerGlyProGlnSerArgArgLys	579		
RESULT 12					
ABB75048					
ID	ABB75048 standard; protein; 586 AA.				
XX	ABB75048;				
XX					
DT	01-MAY-2002 (first entry)				
XX					
DE	Human lung tumour L523S recombinant protein sequence SEQ ID NO:427.				
XX	Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;				
KW	immune response.				
XX	Homo sapiens.				
OS	WO200200174-A2.				
XX					
PD	03-JAN-2002.				
XX					
PF	28-JUN-2001; 2001WO-US021065.				
XX					
PR	28-JUN-2000; 2000US-00606421.				
PR	02-AUG-2000; 2000US-00630940.				
PR	21-AUG-2000; 2000US-00643597.				
PR	15-SEP-2000; 2000US-00662786.				
PR	09-OCT-2000; 2000US-00685696.				
PR	12-DEC-2000; 2000US-00735705.				
PR	07-MAY-2001; 2001US-00850716.				
XX					
PA	(CORI-) CORIXA CORP.				
XX					
PI	Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;				
PI	McNeill PD, Fanger N, Retter MW, Marnierakis M, Fanger GR;				
PI	Vedvick TS, Carter D, Watanabe Y, Peckham DW;				
XX					
DR	WPI; 2002-090513/12.				

Db 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
 QY TCCTCCGCTTGAAGATGATCTTGGAGATTATGATTAAGAGGCTTAAGACACCAAAACG 692
 Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
 QY GCTGACGAGTTCCTCCCTGAAGATCCTGCCCATATAACTTTGTAGGGCGTCTCATGGC 752
 Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
 QY AAGAAGACGACGACCTCAAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAG 812
 Db 294 LysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
 QY TCGTTTGAAGACCTTACCTTTTACAACTTGAAGAGGACCATCTGTGAAGGGGCCATC 872
 Db 314 ProLeuGlnGluLeuThrLeuTyraAsnProGluArgThrIleThrValLysGlyAsnVal 333
 QY GAGAAATTTGTCAGGGCCGAGCAGGAATAATATGAAGAAGTTCCGGAGGCGCTATGAGAA 932
 Db 334 GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
 QY GATGGTGGTGCATGAGC-----TCTCACCTGATCCCTGGCTCAACCTGCTGCTGCTA 986
 Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeu 373
 QY GGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCGC-----CCTCCAGCAGCGCTT 1037
 Db 374 GlyLeuPheProProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
 QY ACTGGGGTGTCTCCCTATGCTTATGCTGCTCCGAGCAGGAGGAGGAGGAGGAGGAGGAG 1097
 Db 394 Thr-----ProProTyProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
 QY TTTATCCCGCCAGGAGTGGCGCCATCATCGCAAGAGGGCGCAGACATCAACAG 1157
 Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGln 430
 QY CTCCTCCGCTTTCAGCGCTCCATCAAGATTGACACACCGCAACACCTGACTCCAAA 1217
 Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
 QY GTTCGTATGTTTATCATCTGACCGCGCAGAGGCGCCAAATCAAGGCTCAGGAGAAATC 1277
 Db 451 ValArgMetValIleIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
 QY TATGGCAACTCAAGAGGAGAACTTCTTGTTCACAGGAGGAGGAGGAGGAGGAGGAGGAG 1337
 Db 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
 QY CACATACGTGTCCAGCAGCAGCTGCGCGGCTCATTCGCAAGAGGTGGAAAAACGGTG 1397
 Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
 QY AACGAGTTCAGAAATTCAGCAGCTGAGGTGGTAGTACCAAGAGACAGACCCCTGAT 1457
 Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
 QY GAGACGACGAGTCTCTGTAATATCATCGACATTTCTATCCAGTCCAGATGCTCAA 1517
 Db 531 GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
 QY CGGAAGATCCGAGACATCTCGGCCAGGTGTAAGCAG---CAGCATCAGAAAGGGA---CAG 1571
 Db 551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnLysAlaLeuGln 570
 QY AGTAACGAGCCGAGCAGCAGGAGAG 1598
 Db 571 SerGlyProProGlnSerArgArgLys 579

ID ADA28438 standard; protein; 579 AA.
 AC ADA28438;
 DT 20-NOV-2003 (first entry)
 XX Human lung tumour protein L5238.
 DE cancer; lung cancer; gene therapy; vaccine; human;
 KW lung squamous cell carcinoma.
 OS Homo sapiens.
 XX US2003064947-A1.
 PD 03-APR-2003.
 PF 30-NOV-2001; 2001US-00007700.
 PR 18-MAR-1998; 98US-00040802.
 PR 27-JUL-1998; 98US-0023912.
 PR 22-DEC-1998; 98US-00221107.
 PR 02-APR-1999; 99US-00285479.
 PR 17-DEC-1999; 99US-00466396.
 PR 30-DEC-1999; 99US-00476496.
 PR 10-JAN-2000; 2000US-00480884.
 PR 22-FEB-2000; 2000US-00510376.
 PR 04-APR-2000; 2000US-00542615.
 PR 28-JUN-2000; 2000US-00606421.
 PR 02-AUG-2000; 2000US-00630940.
 PR 21-AUG-2000; 2000US-00643597.
 PR 15-SEP-2000; 2000US-00662786.
 PR 09-OCT-2000; 2000US-00685696.
 PR 12-DEC-2000; 2000US-00735705.
 PR 07-MAY-2001; 2001US-00850716.
 PR 28-JUN-2001; 2001US-00897778.
 XX (CORI-) CORIXA CORP.
 PA Wang T, Wang A, Skeiky YAM, Li SX, Kalos MD, Henderson RA;
 PI McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
 PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TW;
 XX WPI; 2003-540798/51.
 DR N-PSDB; ADA28437.
 XX New isolated polynucleotides and polypeptides useful for diagnosing,
 PT preventing and/or treating cancer, particularly lung cancer.
 XX Example 2; Page 249-251; 296pp; English.
 CC The invention describes isolated polynucleotides and polypeptides useful
 CC for diagnosing, preventing and/or treating cancer, particularly lung
 CC cancer. A new isolated polynucleotide comprises: any of the 22 fully
 CC defined nucleotide sequences (e.g. 1012, 900 or 2773 bp) given in the
 CC specification; complements of the nucleotide sequences cited above; at
 CC least 10 contiguous residues of the nucleotide sequences cited above; a
 CC sequence that hybridise to any of the nucleotide sequences under highly
 CC stringent conditions; a sequence that is at least 75 or 90% identical to
 CC the above nucleotide sequences; or degenerate variants of the above
 CC nucleotide sequences. The composition and methods are useful in
 CC diagnosing, preventing and/or treating cancer, particularly lung cancer,
 CC in gene therapy and in vaccines. This is the amino acid sequence encoded
 CC by a human lung tumour cDNA isolated from a lung squamous cell carcinoma
 CC that may be useful in the diagnosis and treatment of lung cancer and
 CC other disorders.
 XX SQ Sequence 579 AA;
 Alignment Scores: 6.17e-160 Length: 579
 Pred. No.: 1875.00 Matches: 380
 Score: 84.68% Conservative: 51
 Percent Similarity:

ADA28536
ID ADA28536 standard; protein; 579 AA.
XX
AC ADA28536;
XX
DT 20-NOV-2003 (first entry)
XX
DE Recombinant human lung tumour protein L523S #1.
XX
KW cancer; lung cancer; gene therapy; vaccine; human;
KW lung squamous cell carcinoma.
XX
OS Homo sapiens.
XX
PN US2003064947-A1.
XX
PD 03-APR-2003.
XX
PF 30-NOV-2001; 2001US-00007700.
XX
PR 18-MAR-1998; 98US-00040802.
PR 27-JUL-1998; 98US-00123912.
PR 22-DEC-1998; 98US-00221107.
PR 02-APR-1999; 99US-00285479.
PR 17-DEC-1999; 99US-00466396.
PR 30-DEC-1999; 99US-00476496.
PR 10-JAN-2000; 2000US-00480884.
PR 22-FEB-2000; 2000US-00510376.
PR 04-APR-2000; 2000US-00542615.
PR 28-JUN-2000; 2000US-00606421.
PR 02-AUG-2000; 2000US-00630940.
PR 21-AUG-2000; 2000US-00643597.
PR 15-SEP-2000; 2000US-00662786.
PR 09-OCT-2000; 2000US-00685696.
PR 12-DEC-2000; 2000US-00735705.
PR 07-MAY-2001; 2001US-00850716.
PR 28-JUN-2001; 2001US-00897778.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI McNeill PD, Fanger N, Rector MW, Durham M, Fanger GR, Vedvick TS;
PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TW;
XX
XX WPI; 2003-540798/51.
XX
PT New isolated polynucleotides and polypeptides useful for diagnosing,
PT preventing and/or treating cancer, particularly lung cancer.
XX
PS Claim 9; Page 285-287; 296pp; English.
XX
CC The invention describes isolated polynucleotides and polypeptides useful
CC for diagnosing, preventing and/or treating cancer, particularly lung
CC cancer. A new isolated polynucleotide comprises: any of the 22 fully
CC defined nucleotide sequences (e.g. 1012, 900 or 2773 bp) given in the
CC specification; complements of the nucleotide sequences cited above; at
CC least 10 contiguous residues of the nucleotide sequences cited above; a
CC sequence that hybridize to any of the nucleotide sequences under highly
CC stringent conditions; a sequence that is at least 75 or 90% identical to
CC the above nucleotide sequences; or degenerate variants of the above
CC nucleotide sequences. The composition and methods are useful in
CC diagnosing, preventing and/or treating cancer, particularly lung cancer,
CC in gene therapy and in vaccines. This is the amino acid sequence of a
CC recombinant human lung tumour associated protein.
XX
SQ Sequence 579 AA;

Alignment Scores:		
Pred. No.:	6.17e-160	Length: 579
Score:	1875.00	Matches: 380
Percent Similarity:	84.68%	Conservative: 51
Best Local Similarity:	74.66%	Mismatches: 62
Query Match:	53.10%	Indels: 16

DB:	7	Gaps:	9
US-09-270-437D-7 (1-1946) x ADA28536 (1-579)			
QY	102	CGGTTAGGAGCCGAAATTCAAATCCGAATATTCACCCAGCTCCGATGGGAAGTA	161
DB	77	ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluVal	96
QY	162	CTGACAGCCTGCTGGCTCAGTATGCTACAGTACAGAACTGTGACCACTGACCAACCGAG	221
DB	97	LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp	116
QY	222	AGTCAGAGCGGAGTGGTGAATGTCCATTCCTATTCACCGGAGCAGCAGGCAAGCCATC	281
DB	117	SerGluThrAlaValValAsnValThrTyrSerIysAspGlnAlaArgGlnAlaLeu	136
QY	282	ATGAAGCTGAATGGCCACCACTGTGGAGAACCACTGCCCTGAAGGTCTCTACATCCCGAT	341
DB	137	AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp	156
QY	342	GAGCAGATAGCA---CAGGGACCTGAGAAATGGGCGCCGAGG-----GGCTTTGGCTCT	392
DB	157	GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGln	176
QY	393	CGGGCTCAGCCCGCCAGGGCTCACCTCTGGCAGCGGGGCCCCAGCCAGCAGCAGCAA	452
DB	177	ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro	193
QY	453	GTGACATCCCTTCGGCTCTCTGTGCCCCACCACTATGTGGTGCCATTATGGCAAG	512
DB	194	CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys	213
QY	513	GAGGGGCCACATCCGCAACATCAAAACAGACCCAGTCCAAATAGACATGATGATAGG	572
DB	214	GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg	233
QY	573	AAGGAGACGCGAGGTGACAGTGTGAAAGCCATCTGCTGCATCCACCCCTCAGGGCTGC	632
DB	234	LysGluAsnAlaGlyAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr	253
QY	633	TCCTCCGCTGTGAAGATCTTGAGATTATGCAATAAGAGCTTAAGACACCAAAACG	692
DB	254	SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe	273
QY	693	GCTGACGAGTTCCCTGAAAGATCTTGGCCCAATAACTTTGTAGGGCGTCTCATGGC	752
DB	274	ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly	293
QY	753	AAGGAGGACGAACTGAAGAGGTAGAGCAAGATACCGAGACAAATACCATCTCC	812
DB	294	LysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLysIleThrIleSer	313
QY	813	TCGTTGCAAGACCTTTACCCCTTTACAAACCTGAGAGACCATCATCTGTGAAGGGGCGATC	872
DB	314	ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal	333
QY	873	GAGAAATGTTGCGGCGCAGCAGGAATAATATGAGAAAGTTTCGGGAGCCCTATGAGAT	932
DB	334	GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyrGluAsn	353
QY	933	GATGTGGTCCCATGAGC-----TCTACCTGATCCCTGGCTGAACCTGGCTGCTGTA	986
DB	354	AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeu	373
QY	987	GGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCGC-----CCTCCAGCAGCGTT	1037
DB	374	GlyLeuPheProProThrSerGlyMetProProProThrSerGlyProProSerAlaMet	393
QY	1038	ACTGGGGCTGCTCCCTATAGCTCTTTATGCGAGGCTCCCGCAGCAGAGATGGTCAGGTG	1097
DB	394	Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu	410
QY	1098	TTTATCCCGCCGAGGCGAGTGGGCGCATCATCGGCAAGAGAGGGGCGAGCAGCATCAAACAG	1157

PI Moneill PD, Fanger N, Retter WM, Durham M, Fanger GR, Vedwick TS;
 PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
 XX WPI; 2002-583465/62.
 DR N-PSDB; ABQ92483.
 XX Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
 PT the polynucleotides, useful in pharmaceutical compositions such as
 PT vaccines and as markers to indicate the presence of lung cancer.
 XX
 XX Claim 9; Page 372-374; 381pp; English.

CC The present invention describes isolated human lung carcinoma
 CC polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
 CC activity, and can be used in gene therapy and in vaccines. Compositions
 CC comprising (I) or (II) can be used for stimulating an immune response in
 CC a patient and for treating lung cancer in a patient. Oligonucleotides of
 CC (I) can be used for detecting the presence of a cancer in a patient, by
 CC obtaining a biological sample from the patient, contacting the biological
 CC sample with the oligonucleotide, detecting in the sample, an amount of
 CC polynucleotide that hybridises to the oligonucleotide and comparing the
 CC amount of polynucleotide that hybridises to the oligonucleotide to a
 CC predetermined cut-off value, and determining the presence of a cancer in
 CC the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
 CC vaccines. (I) is useful as a marker to indicate the presence or absence
 CC of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP61866 to
 CC ABP61992 represent sequences used in the exemplification of the present
 CC invention

XX SQ Sequence 579 AA;

Alignment Scores:

Pred. No.: 6.17e-160 Length: 579
 Score: 1875.00 Matches: 380
 Percent Similarity: 84.68% Conservative: 51
 Best Local Similarity: 74.66% Mismatches: 62
 Query Match: 53.10% Indels: 16
 DB: 5 Gaps: 9

US-09-270-437D-7 (1-1946) x ABP61973 (1-579)

QY 102 CGGGTGAAGACCGCGAAATTCGAATATCCACCCAGCTCCGATGGGAAGTA 161
 DB 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluVal 96
 QY 162 CTGACAGCTGCTGGCTAGTATGTGTACAGTAGAAGCTGTGAGCAAGTGAACACCGAG 221
 DB 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
 QY 222 AGTGAGACGCGAGTGTGAATGTACCTATTCGAACCGGAGACGACGACGACCATC 281
 DB 117 SerGluThrAlaValAlaValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
 QY 282 ATGAAGCTGAATGCCACCGCTGGAGAACCATCCCTGAAGGTCTCTACATCCCGCAT 341
 DB 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
 QY 342 GAGCAGATAGCA---CAGGGACCTGAGAAATGGGCGCCGAGGG-----GGCTTTGGCTCT 392
 DB 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGln 176
 QY 393 CGGGGTGACGCCCGCCAGGGCTACCTGTGGCAGCGGGGGCCCCCAGCCAGCAGCAAA 452
 DB 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
 QY 453 GTGACATCCCGCTTCGGCTCTCGGTGCCACCCAGTATGGGTGCGCATATTGGCAAG 512
 DB 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
 QY 513 GAGGGGCGCCACCATCCGCAACATCACAACAAACAGACCCAGTCCAAAGATAGCGTATAGG 572
 DB 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrLysSerLysIleAspValHisArg 233

QY 573 AAGAGAACCCAGGTGCAGCTGAAAGAACCCATCAGTGTGCACCTCCACCCCTGAGGGTGC 632
 DB 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
 QY 633 TCCTCCGCTTGTAGATGATCTTGGAGATTATCATAAAGAGGCTAAGGACACCAACG 692
 DB 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
 QY 693 GCTGACGAGGTTCCTCTGAAGATCTCGGCCATAATAACTTTGTAGGCGCTCTCATGGC 752
 DB 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
 QY 753 AAGAGAGCGAAGACCTGAAGAGGTAGACAGATACCGAGACAAAATACCATCTCC 812
 DB 294 LysGluGlyArgAsnLeuLysLysIleGluGluAspThrAspThrLysIleThrIleSer 313
 QY 813 TCGTGTCAACACCTTACCTTTACACCTGAGAGACCATCTCTGTAAGGGGCGCATC 872
 DB 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
 QY 873 GAGAATTGTTCAGGGCGGAGCAAGAAATATCAAGAAAGTTCGGGAGGCTATGAGAT 932
 DB 334 GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
 QY 933 GATGTGGCTCCCATGACC-----TCTCACCTGATCCCTGGCTGACCTGGCTGCTGTA 986
 DB 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeu 373
 QY 987 GGTCTTTTCCAGCTTCATCCAGCGAGTCCCGCG-----CTCCACGACGAGT 1037
 DB 374 GlyLeuPheProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
 QY 1038 ACTGGGCTGCTCCCTATAGCTCTTTATGAGGCTCCCGAGGAGATGGTGAGGTG 1097
 DB 394 Thr-----ProProTyrProGlnPheGluInSer---GluThrGluThrValHisLeu 410
 QY 1098 TTTATCCCGCCAGGACGTGGGCGCATCATCGCAAGAGGGGCGACGACATCAACAG 1157
 DB 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGln 430
 QY 1158 CTCTCCCGTTCCTCCAGGCTCCATCAAGATTGCACCCCGAAACCTCTGACTCCAAA 1217
 DB 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
 QY 1218 GTTCGTATGTTATCATCATCTGACCGCCAGAGCCCAATTCAGGCTCAGGGAAGATC 1277
 DB 451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
 QY 1278 TATGGCAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGAGTGAAGCTGGAGACC 1337
 DB 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
 QY 1338 CACATACGTGTCCAGCTCAGCAGCTGGCGGGCTCATTTGGCAAAAGTGGAAAAAGCGTG 1397
 DB 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyGlyLysThrVal 510
 QY 1398 AACGAGTTGCAATTTGACGCGAGCTGAGGTGTAGTACCAAGACACAGACCCCTGAT 1457
 DB 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
 QY 1458 GAGAACGACGAGTCTCGTGAATAATCATCGGACATTTCTATCCAGTCAAGATGGCTCAA 1517
 DB 531 GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
 QY 1518 CGGAAGATCCGACATCTCTGGCCCGAGTTAAGCAG---CAGCATCAGAGGGA---CAG 1571
 DB 551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnLysAlaLeuGln 570
 QY 1572 AGTAACGAGCCCGACGACGAGGAAG 1598
 DB 571 SerGlyProGlnSerArgArgLys 579

RESULT 9

Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
QY 693 GCTCAGCAGGTTCCCTGAGATCCTGGCCCATATTAATCTTGTAGGGCGTCTATTGGC 752
Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY 753 AAGGAAGCAGCAACCTCGAAGAGGTAGAGCAATACCGAGACAAATAATCACCATTCC 812
Db 294 LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
QY 813 TCGTTGAGACCTTACCTTTACACCTGAGAGGACCATCACTGTGAGGGGCCATC 872
Db 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY 873 GAGAAATTTGTCAGGGCCGAGCAAGAAATATGAAGAAAGTTCGGGAGGCTTATCAGAA 932
Db 334 GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
QY 933 GATGTGGTGCATGAGC-----TCTCACTGATCCCTGGCCTGAACCTGGCTCTCTA 986
Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeu 373
QY 987 GGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCTCCGAGCAGCGTT 1037
Db 374 GlyLeuPheProProThrSerGlyMetProProThrSerGlyProProSerAlaMet 393
QY 1038 ACTGGGGTGTCTCCTATAGCTCTTTATGAGGCTCCGAGCAGAGATGTCAGGTG 1097
Db 394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
QY 1098 TTTATCCCGCCGAGCAGTGGCGCCATCATCGCAAGAAGGGCGACACATCAACAG 1157
Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyHisIleLysGln 430
QY 1158 CTCTCCCGTTTGGCAGCGCTCCATCAGATTGCAACCCGCAACCTGACCTCCAA 1217
Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY 1218 GTTCTGTATGTTATCATCTGACCGCGCAGAGGCCCAATTCAGAGCTCAGGGAAGATC 1277
Db 451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY 1278 TATGCAAACTCAAGGAGGAGAACTTCTTTGTCCCAAGGAGGAAGTGAAGCTGAGACC 1337
Db 471 TyrGlyLysIleLysGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY 1338 CACATACGTGTCCAGCATCAGCAGCTGCGCGGTTCATTGGCAAGGTGGAAAAACGGTG 1397
Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
QY 1398 AACGAGTTGCAGAAATTCAGGCGAGCTGAGGTGGTAGTACCAAGAGACAGACCCCTGAT 1457
Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
QY 1458 GAGAACGACGAGTTCATCGTGAATATCATCGACATTTTATGCCAGTCAGATGCTCAA 1517
Db 531 GluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
QY 1518 CGAAGATCCGAGACATCTGTGCCCGCAGGTATAGCAG---CAGCATCAGAGGGA---CAG 1571
Db 551 ArgLysIleGlnGluIleThrGlnValLysGlnHisGlnGlnLysAlaLeuGln 570
QY 1572 ACTAACCGCGCCGAGCAGGAGGAAG 1598
Db 571 SerGlyProGlnSerArgArgLys 579
RESULT 6
ABP61917
ID ABP61917 standard; protein; 579 AA.
XX
AC ABP61917;
XX

DT 07-OCT-2002 (first entry)
DE Human lung cancer associated protein sequence SEQ ID NO:348.
KW Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.
OS Homo sapiens.
PN WO200247534-A2.
PD 20-JUN-2002.
XX 30-NOV-2001; 2001WO-US047576.
PF 12-DEC-2000; 2000US-00735705.
PR 07-MAY-2001; 2001US-00850716.
PR 28-JUN-2001; 2001US-00897778.
XX (CORI-) CORIXA CORP.
XX Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX WPI; 2002-583465/62.
DR N-PSDB; ABQ92440.
XX Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
PT the polynucleotides, useful in pharmaceutical compositions such as
PT vaccines and as markers to indicate the presence of lung cancer.
XX
PS Example 2; Page 337-339; 381pp; English.
CC The present invention describes isolated human lung carcinoma
CC polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
CC activity, and can be used in gene therapy and in vaccines. Compositions
CC comprising (I) or (II) can be used for stimulating an immune response in
CC a patient and for treating lung cancer in a patient. Oligonucleotides of
CC (I) can be used for detecting the presence of a cancer in a patient, by
CC obtaining a biological sample from the patient, contacting the biological
CC sample with the oligonucleotide, detecting in the sample, an amount of
CC polynucleotide that hybridises to the oligonucleotide and comparing the
CC amount of polynucleotide that hybridises to the oligonucleotide to a
CC predetermined cut-off value, and determining the presence of a cancer in
CC the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
CC vaccines. (I) is useful as a marker to indicate the presence or absence
CC of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP61866 to
CC ABP61992 represent sequences used in the exemplification of the present
CC invention
XX SQ Sequence 579 AA;
Alignment Scores:
Pred. No.: 6,17e-160 Length: 579
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best Local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 5 Gaps: 9
US-09-270-437D-7 (1-1946) x ABP61917 (1-579)
QY 102 CCGGTTAGAGCGCGAAATTCGAATATTCACCCAGCTCCGATGGGAGTA 161
Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuTrpGluVal 96
QY 162 CTGGACAGCTGCTGCTCAGTATGTACAGTAGAGAAGTGTGAGCAAGTGAACACCGAG 221
Db 97 LeuAspSerLeuLeuValGlnIleGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGAGACGGCAGTGTGAATGTCACTTATTCACACCGGAGCAGCAGGCAAGCATC 281
Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136

ID ABB75053 standard; protein; 579 AA.
 AC ABB75053;
 DT 01-MAY-2002 (first entry)
 XX Human lung tumour L523S recombinant protein sequence SEQ ID NO:446.
 DE Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
 KW immune response.
 KW Homo sapiens.
 OS
 XX W0200200174-A2.
 XX 03-JAN-2002.
 XX 28-JUN-2001; 2001WO-US021065.
 XX 28-JUN-2000; 2000US-00606421.
 PR 02-AUG-2000; 2000US-00630940.
 PR 21-AUG-2000; 2000US-00643597.
 PR 15-SEP-2000; 2000US-00662786.
 PR 09-OCT-2000; 2000US-00685696.
 PR 12-DEC-2000; 2000US-00735705.
 PR 07-MAY-2001; 2001US-00850716.
 XX (CORI-) CORIXA CORP.
 PA Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
 PI McNeill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;
 PI Vedvick TS, Carter D, Watanabe Y, Peckham DW;
 XX WPI; 2002-090513/12.
 DR N-PSDB; ABL49297.
 XX Polynucleotides encoding lung tumor polypeptides, useful for treating
 PT lung cancer or stimulating an immune response.
 PT Claim 2; Page 365-367; 374pp; English.
 PS The present invention describes human lung tumour proteins. Human lung
 CC tumour proteins and polynucleotides have cytostatic and immunostimulant
 CC activities, and can be used in vaccine production. Compositions
 CC comprising the lung tumour proteins, polynucleotides, antibodies, fusion
 CC proteins, T cell populations, or antigen presenting cells that express
 CC the lung tumour proteins are useful for treating lung cancer or
 CC stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
 CC ABB75070 represent sequences used in the exemplification of the present
 CC invention
 XX Sequence 579 AA;
 SQ
 Alignment Scores:
 Pred. No.: 6.17e-160 Length: 579
 Score: 1875.00 Matches: 380
 Percent Similarity: 84.68% Conservative: 51
 Best Local Similarity: 74.66% Mismatches: 62
 Query Match: 53.10% Indels: 16
 DB: 5 Gaps: 9
 US-09-270-437D-7 (1-1946) x ABB75053 (1-579)
 QY 102 CGGGTAGGACGGGAAATCCAAATCCGAAATATCCACCCCGAGCTCCGATGGAGTA 161
 DB 77 ArgGlnArgIleArgIysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluVal 96
 QY 162 CTGACAGCTCGTGGCTCAGTATGTACAGTAGAGAACTGTGAGCAAGTGAACCCGAG 221
 DB 97 LeuAspSerLeuLeuValGlnTrpGlyValValGluSerCysGluGlnValAsnThrAsp 116
 QY 222 AGTGAGACGGCAGTGGTGAATGTCACTATTCCAAACCGGAGGAGACACGAGGCATC 281
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
 QY 282 ATGAAGCTGAATGGCCACCAAGTTGGAGAACCATGCCCTGAAGGTCTCTCATATCCCCGAT 341
 Db 137 AsplysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
 QY 342 GAGCAGATAGCA---CAGGACCTGAGAAATGGGGCCGAGGG-----GGCTTTGGTCT 392
 Db 157 GluThrAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGln 176
 QY 393 CGGGTCCAGCCCGCCAGGCTCACCTGTGGCAGCGGGGGCCCGAGCAGACGACGACAA 452
 Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
 QY 453 GTGGACATCCCTTCCGCTCTCTGGTCCCAACCCAGTATCTGGTGCATATTATGGCAAG 512
 Db 194 CysAspLeuProLeuArgLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
 QY 513 GAGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCAGATAGACGTGCATAGG 572
 Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
 QY 573 AAGGAGAACGACGCTGCAAGGAAAGCCATCAGTGTGCACCTCCACCCCTGAGGGCTGC 632
 Db 234 LysGluAsnAlaGlyAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
 QY 633 TCCTCCGCTTGAAGATGATCTTGGAGATTATCATAAAGAGCTAAGGACACCAAAACG 692
 Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
 QY 693 GCTGACGAGTTCCCTCGAAGATCTTGGGCCCAATAAATCTTGTAGGCGCTCTCATGGC 752
 Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
 QY 753 AAGGAAGGACGGAACCTGAAAGAGTAGAGCAAGATACCGAGACAAAATCACCATCTCC 812
 Db 294 LysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
 QY 813 TCCTTGAAGACCTTACCTTTTACAACCTTGAGAGGACCATCATCTGTGAAGGGGGCCATC 872
 Db 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
 QY 873 GAGAAATGTTCAGGGCCGAGCAGGAAATTAATCAAGAAATTCGGAGGCGCTATGAGAT 932
 Db 334 GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
 QY 933 GATGTGGCTGCCATGAGC-----TCTCACCTCATCCCTGGCTGGAACCTGGCTGTGTA 986
 Db 354 AsplleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeu 373
 QY 987 GGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CTCCACGACGCGTT 1037
 Db 374 GlyLeuPheProProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
 QY 1038 ACTGGGGCTGCTCCCTATAGCTCTTTATGTCAGGCTCCGAGCAGGAGATGGTGCAGGTG 1097
 Db 394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
 QY 1098 TTTATCCCGCCCGCAGGAGTGGGGCGGCATCATCGCAAGAGGGGCGAGCACATCAACAG 1157
 Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGln 430
 QY 1158 CTCTCCGGTGTTCGAGCGCTCCATCAAGATTGACCCACCCGAAACACTGACTCCAA 1217
 Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
 QY 1218 GTTCGTATGTTATCATCTGACCGCCGAGGGCCCAATTCAGGCTCAGGGAAGATC 1277
 Db 451 ValArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 470
 QY 1278 TATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACC 1337
 Db 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490

Example; Fig 1A-D; 79pp; English.

The present sequence represents a murine c-myc coding region determinant binding protein (CRD-BP). The presence or absence of a tumor can be determined by determining the levels of CRD-BP present in the suspect tissue, where the CRD-BP shields c-myc RNA from endonucleolytic attack and so prolongs its half-life. The methods are used for diagnosing the presence or absence of a tumor in a human, especially breast, colon and pancreatic cancer. They are also used to inhibit cancer cell growth.

Sequence 577 AA;

Alignment Scores:

Pred. No.:	1.5e-215	Length:	577
Score:	2491.00	Matches:	493
Percent Similarity:	99.00%	Conservative:	1
Best Local Similarity:	98.80%	Mismatches:	3
Query Match:	70.55%	Indels:	2
DB:	2	Gaps:	1

US-09-270-437D-7 (1-1946) x AAY30649 (1-577)

Qy	108	AGGAGCCGGAAAAATTCAAAATCCGAAATATTCACCCCAAGCTCCGATGGGAAGTACTCTGAC	167
Db	79	ArgSerArgLysIleGlnIleArgAsnIleProProGlnLeuArgTrpGluValLeuAsp	98
Qy	168	AGCCTGCTGGCTCAGTATGGTACAGTAGAAGACTGTGAGCAAGTGAACACCCAGAGTGG	227
Db	99	SerLeuLeuAlaGlnTyrGlyThrValGluAsnCysGluGlnValAsnThrGluSerGlu	118
Qy	228	ACGGCAGTGTGAAATGTCACTATTCACCAACGGGAGCAGACCAGCAAGCCATCATGAAG	287
Db	119	ThrAlaValValAsnValThrTyrSerAsnArgGluGlnThrArgGlnAlaIleMetLys	138
Qy	288	CTGAATGGCCACCGATTGGAGAACCATGCCCTGAAGTCTCTCTACATCCCCGATCAGCAG	347
Db	139	LeuAsnGlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyrIleProAspGluGln	158
Qy	348	ATAGCAGAGGACCTTGACAATGGGCGCCGAGGGGGCTTTGGCTCTCGGGGTGAGCCCGC	407
Db	159	IleThrGlnGlyProGluAsnGlyArgArgGlyGlyPheGlySerArgGlyGlnProArg	178
Qy	408	CAGGGCTCACCTGTGGCAGCGGGGGCCCCAGCCAGCAGCAGCAAGTGACATCCCCCTT	467
Db	179	GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnProValAspIleProLeu	198
Qy	468	CGGCTCTGTGTGCCACCCAGTATGTGGGTGCCATTATGGCAAGAGGGGGCCACCATC	527
Db	199	ArgLeuLeuValProThrGlnTyrValGlyAlaIleIleGlyLysGluGlyAlaThrIle	218
Qy	528	CGCAACATCACAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAAGGAGAACCGAGT	587
Db	219	ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGly	238
Qy	588	GCAGCTGAAAAAGCCATCAGTGTGCACATCCACCCCTGAGGGTGCTCTCTCCGCTTGTAA	647
Db	239	AlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCysSerSerAlaCysLys	258
Qy	648	ATGATCTTGGAGATTATGCATAAGAGGCTAAGACACAAAACGGCTGACGAGGTTCC	707
Db	259	MetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThrAlaAspGluValPro	278
Qy	708	CTGAAGATCCTGGGCCATAATAACTTTCTPAGGGGCTCATTTGGCAAGGAGGACGGAC	767
Db	279	LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn	298
Qy	768	CTGAAGAAGGTAGACCAAGATACCAGACAATAATCACCATCTCTCTCTGTCGAACACCTT	827
Db	299	LeuLysLysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeu	318
Qy	828	ACCCCTTTACACCTGAGAGACCATCATCTGTGAAGGGGGCCATCGAATTTGTCAGG	887
Db	319	ThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIleGluAsnCysCysArg	338

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 23, 2004, 10:51:25 ; Search time 85.4096 Seconds
(without alignments)
12875.302 Million cell updates/sec

Title: US-09-270-437D-7
Perfect score: 3531
Sequence: 1 gctgtagggagggtggg.....atttcctcaggttttaaaa 1946

Scoring table: BLOSUM62
Ygapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US0270437/runat_23072004_095738_27403/app.query.fasta_1.5582
-DB=A Genesep29Jan04 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0270437 @CGN 1 1 185 @runat_23072004_095738_27403 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

- Database :
- 1: Genesep29Jan04:*
 - 2: Genesep1980s:*
 - 3: Genesep1990s:*
 - 4: Genesep2000s:*
 - 5: Genesep2001s:*
 - 6: Genesep2002s:*
 - 7: Genesep2003as:*
 - 8: Genesep2003bs:*
 - 9: Genesep2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2491	70.5	577	2	Aay30649 A murine
2	1875	53.1	579	3	Aab11365 Human lun
3	1875	53.1	579	5	Abb75053 Human lun
4	1875	53.1	579	5	Abb74997 Human lun
5	1875	53.1	579	5	Abb75054 Human lun
6	1875	53.1	579	5	Abp61917 Human lun
7	1875	53.1	579	5	Abp61974 Human lun
8	1875	53.1	579	5	Abp61973 Human lun
9	1875	53.1	579	7	Ada28536 Recombina
10	1875	53.1	579	7	Ada28539 Recombina

11	1875	53.1	579	7	ADA28438	Ada28438 Human lun
12	1875	53.1	586	5	ABP75048	Abp75048 Human lun
13	1875	53.1	586	5	ABP61968	Abp61968 Human lun
14	1875	53.1	586	7	ADA28517	Ada28517 Recombina
15	1873	53.0	579	3	AAB11328	Aab11328 Human lun
16	1873	53.0	579	5	ABB74960	Abb74960 Human lun
17	1873	53.0	579	5	ABP61880	Abp61880 Human lun
18	1873	53.0	579	5	ABP61960	Abp61960 Human lun
19	1873	53.0	579	7	ADA28266	Ada28266 Human lun
20	1873	53.0	579	7	ADD14066	Add14066 Human src
21	1802	51.0	619	4	ABG21963	Abg21963 Novel hum
22	1763.5	49.9	614	4	ABG06794	Abg06794 Novel hum
23	1763.5	49.9	620	4	AAU16163	Aau16163 Human nov
24	1763.5	49.9	620	6	ABU55232	Abu55232 Human nov
25	1740	49.3	583	4	ABG12592	Abg12592 Novel hum
26	1697	48.1	556	5	ABG96346	Abg96346 Human ova
27	1697	48.1	594	4	ABG06795	Abg06795 Novel hum
28	1585	44.9	319	4	AAM93826	Aam93826 Human pol
29	1573.5	44.6	555	6	ABU89799	Abu89799 Novel hum
30	1258	35.6	261	4	AAU16161	Aau16161 Human nov
31	1258	35.6	261	6	ABU55230	Abu55230 Human nov
32	1228	34.8	250	4	AAU16579	Aau16579 Human nov
33	1228	34.8	250	6	ABU55648	Abu55648 Human nov
34	863	24.4	558	4	ABB58367	Abb58367 Drosophil
35	653	18.5	171	4	AAU16166	Aau16166 Human nov
36	605.5	17.1	187	6	ABU55235	Abu55235 Human nov
37	605.5	17.1	187	6	AAO23971	Aao23971 Human IGF
38	510	14.4	192	6	AAU16164	Aau16164 Human nov
39	510	14.4	192	6	ABU55233	Abu55233 Human nov
40	502	14.2	171	4	AAU16583	Aau16583 Human nov
41	502	14.2	171	6	ABU55652	Abu55652 Human nov
42	465.5	13.2	148	4	ABG21962	Abg21962 Novel hum
43	410	11.6	266	4	ABG12593	Abg12593 Novel hum
44	409.5	11.6	209	4	ABG21961	Abg21961 Novel hum
45	292	8.3	93	4	AAM38501	Aam38501 Peptide #

ALIGNMENTS

RESULT 1
AAY30649
ID AAY30649 standard; protein; 577 AA.
XX
AC AAY30649;
XX
DT 17-NOV-1999 (first entry)
XX
DE A murine c-myc coding region determinant binding protein.
XX
DE c-myc coding region determinant binding protein; CRD-BP; tumor; c-myc;
XX
DE endonucleolytic attack; half-life; breast cancer; colon cancer;
XX
DE pancreatic cancer.
XX
OS Mus musculus.
XX
FN WO9946594-A2.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004897.
XX
PR 09-MAR-1998; 98US-0077372P.
XX
PR (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PA Ross J;
XX
PI WPI; 1999-551506/46.
XX
DR N-PSDB; AA210617.
XX
PT Diagnosing presence or absence of a tumor in a human by examining c-myc
XX
PT coding region determinant-binding protein.

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4399556"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 88"
/notes="Organ: small intestine; Vector: pCMV-SPORT6;
Site 1: Not1; Site 2: SalI; Cloned unidirectionally;
oligo-dT primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH MGC library."
```

ORIGIN

Query Match	34.0%	Score 662;	DB 10;	Length 896;
Best Local Similarity	93.4%	Pred. No. 1.8e-156;		
Matches 727;	Conservative 0;	Mismatches 40;	Indels 11;	Gaps 3;

QY 821 AGACCTTACCTTTTACAACCTTGAGAGGACCATCACTGTGAAGGGGCCATCGAGAATTG 880
Dd |||||||
1 AGACCTTACCTTTTACAACCTTGAGAGGACCATCACTGTGAAGGGGCCATCGAGAATTG 60
QY 881 TTGCAGGGCCGACAGAGAAATATGAAGAAAGTTTCGGAGGCCCTATGAGAAATGATGTGGC 940
Dd |||||||
61 TTGCAGGGCCGACAGAGAAATATGAAGAAAGTTTCGGAGGCCCTATGAGAAATGATGTGGC 120
QY 941 TGCCTATGAGC-----TCTCACCTGATCCCTGGCCTGAACCTGGCTGCTGTAGGTCTTTT 994
Dd |||||||
121 TGCCTATGAGCCTGAGTCTCACTGATCCCTGGCCTGAACCTGGCTGCTGTAGGTCTTTT 180
QY 995 CCAGGTTTANTCAGGCGAGTCCCGCGCTCCCGAGCGGTTACTGGGGCTGCTCCCTA 1054
Dd |||||||
181 CCAGGTTTATCCAGGCGAGTCCCGCGCTCCCGAGCGGTTACTGGGGCTGCTCCCTA 240
QY 1055 TAGCTCCTTTATGAGGCTCCCGAGCAGAGATGGTGCAGGTGTTTATCCCGCCCGCAGGC 1114
Dd |||||||
241 TAGCTCCTTTATGAGGCTCCCGAGCAGAGATGGTGCAGGTGTTTATCCCGCCCGCAGGC 300
QY 1115 AGTGGCGCCCATCATCGGCAAGAGGGCGACACATCAACACAGCTCTCCCGGTTTGCAG 1174
Dd |||||||
301 AGTGGCGCCCATCATCGGCAAGAGGGCGACACATCAACACAGCTCTCCCGGTTTGCAG 360
QY 1175 CGCCTCCATCAAGATTGCACACCCGAAACACCTGACTCCAAAGTTGCTATGTTATCAT 1234
Dd |||||||
361 CGCCTCCATCAAGATTGCACACCCGAAACACCTGACTCCAAAGTTGCTATGTTATCAT 420
QY 1235 CACTGGACCGCCAGAGCCCAATTCAAGCTCAGGAGAGATCTATGGCAACTCAAGGA 1294
Dd |||||||
421 CACTGGACCGCCAGAGCCCAATTCAAGCTCAGGAGAGATCTATGGCAAACTCAAGGA 480
QY 1295 GGAGAACTTTTGTGTCCTCAAGGAGGAAGTGAAGCTGGAGACCCACATAGTGTGCCAGC 1354
Dd |||||||
481 GGAGAACTTTTGTGTCCTCAAGGAGGAAGTGAAGCTGGAGACCCACATAGTGTGCCAGC 540
QY 1355 ATCAGCAGCTGGCCGGGTGATTGGCAAAAGTGAAGGCAAAACGGTGAAACGAGTTGCAGAAATTT 1414
Dd |||||||
541 ATCAGCAGCTGGCCGGGTGATTGGCAAAAGTGAAGGCAAAACGGTGAAACGAGTTGCAGAAATTT 600
QY 1415 GACGGCAGCTGAGGTAGTACCRAGAGACAGACCCCTGATGAGAACGACAGGTGAT 1474
Dd |||||||
601 GACGGCAGCTGAGGTAGTACCRAGAGACAGACCCCTGATGAGAACGACAGGTGAT 660
QY 1475 CGTGAAA-----ATCATCGGACATTTCTATGTCAG-TCAGATGGCTCAACGGAGATCCGA 1529
Dd |||||||
661 CGTGAAACATCCATCCGGAACATTTCTATGCCAGCTCAGATGGCTCAGGAGATCCGAG 720
QY 1530 GACATCCTGGCCCGAGGTTAAGCAGCAGCATCAAGAGGGAAGAGTAAACAGGCCCGCAGG 1587
Dd |||||||
721 AACTTCTGGCCCGAGGTTAAGCGCGCAGCATCGAAAGGACGGGTTCCCGCGCGCGG 778

Search completed: July 24, 2004, 16:42:21
Job time : 4881.33 secs

Db	481	TTCCAGTGGAGAACCTTGATCTCTCAGCCCCAACACCCACCAATTTGGCCCCACACTGT	540
Qy	1809	NTGCCCTCCGGGTGTGAGAAATTTAGCGAAGGCACTTTTAAACGTGGATGTTTAAA	1868
Db	541	CTGCCCCCTCGGGGTGTGAGAAATTTCTAGCGAAGGCACTTTTAAACGTGGATGTTTAAA	600
Qy	1869	GAAGCTCTCAGGCCCCCACCAGAGGGTGGATCACACCTCAGTGGGAGAAATAAAT	1928
Db	601	GAAGCTCTCAGGCCCCCACCAGAGGGTGGATCACACCTCAGTGGGAGAAATAAAT	660
Qy	1929	TTCCCTCAGGTTTT	1942
Db	661	TTCCCTCAGGTTTT	674
RESULT 14			
LOCUS	BQ651665	955 bp	mRNA
DEFINITION	AGENCOURT 8297761 NIH_MGC_100 Homo sapiens cdna clone IMAGE:6269842		EST 15-JUL-2002
ACCESSION	BQ651665		
VERSION	BQ651665.1	GI:21775837	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	1 (bases 1 to 955)		
COMMENT	NIH-MGC http://mgi.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: CGAP (Stanford) cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LICM2444 row: a column: 11 High quality sequence stop: 588. Location/Qualifiers		
FEATURES			
source	1..955	/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:6269842"	
		/tissue type="hepatocellular carcinoma, cell line"	
		/lab host="DH10B (phage-resistant)"	
		/clone lib="NIH MGC 100"	
		/note="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."	
ORIGIN			
Query Match	34.0%;	Score 662.2;	DB 13; Length 955;
Best Local Similarity	98.8%;	Pred. No. 1.7e-156;	
Matches 667;	Conservative 0;	Mismatches 8;	Indels 0; Gaps 0;
Qy	264	CAGACAGGCAAGCCATCATGAAGCTGAATGCCACCACTGCGAGAACCATGCCCTGAAG	323
Db	1	CAGACAGGCAAGCCATCATGAAGCTGAATGCCACCACTGCGAGAACCATGCCCTGAAG	60
Qy	324	GTCTCTACATCCCGATGACGATAGACACAGGACCTGAGATGGCGCGCGAGGGGC	383

Db	61	GTCTCTACATCCCGATGACGATAGACACAGGACCTGAGAATGGCGCGCGAGGGGC	120
Qy	384	TTTGGCTCTCGGGTGAAGCCCGCCAGGGCTCACTGTGGCAGCGGGGCCCGCCAGCAAG	443
Db	121	TTTGGCTCTCGGGTGAAGCCCGCCAGGGCTCACTGTGGCAGCGGGGCCCGCCAGCAAG	180
Qy	444	CAGCAGCAAGTGGACATCCCTTCGGCTCTCTGTGTGCCACCCAGTATGTGGGTGCCATT	503
Db	181	CAGCAGCAAGTGGACATCCCTTCGGCTCTCTGTGTGCCACCCAGTATGTGGGTGCCATT	240
Qy	504	ATTGGCAAGAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAAGATAC	563
Db	241	ATTGGCAAGAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAAGATAC	300
Qy	564	GTGCATAGGAAGAGAAACGACAGTGCAGTGAAGAACCATCAGTGTGCATCCACCCCT	623
Db	301	GTGCATAGGAAGAGAAACGACAGTGCAGTGAAGAACCATCAGTGTGCATCCACCCCT	360
Qy	624	GAGGGCTCTCTCCGCTTGTAAAGATGATCTTGAGATTTATGCAATAAGAGGCTAAGAC	683
Db	361	GAGGGCTCTCTCCGCTTGTAAAGATGATCTTGAGATTTATGCAATAAGAGGCTAAGAC	420
Qy	684	ACCAAAACGGCTGACGAGGTTCCCTGAGATCTCTGGCCCAATAATACTTTGTAGGGCT	743
Db	421	ACCAAAACGGCTGACGAGGTTCCCTGAGATCTCTGGCCCAATAATACTTTGTAGGGCT	480
Qy	744	CTCATTGGCAAGAGAGACGAACTGAGAGAGTAGAGCAAGATACCGACACAAAATC	803
Db	481	CTCATTGGCAAGAGAGACGAACTGAGAGAGTAGAGCAAGATACCGACACAAAATC	540
Qy	804	ACCATCTCTCGTTCGAGACCTTACCCCTTACAAACCTTGAGAGACCATCACTGTGAAG	863
Db	541	ACCATCTCTCGTTCGAGACCTTACCCCTTACAAACCTTGAGAGACCATCACTGTGAAG	600
Qy	864	GGGGCCATCGAGAATTGTCAGGGCCGAGCAGGAAATTAATGAAGAAAGTTGGGAGGCC	923
Db	601	GGGGCCATCGAGAATTGTCAGGGCCGAGCAGGAAATTAATGAAGAAAGTTGGGAGGCC	660
Qy	924	TATGAGAAATGATGTG 938	
Db	661	TATGCAATGACGTG 675	
RESULT 15			
LOCUS	BF984962	896 bp	mRNA
DEFINITION	602308484F1 NIH_MGC_88 Homo sapiens cdna clone IMAGE:4399556 5',		linear EST 23-JAN-2001
ACCESSION	BF984962		
VERSION	BF984962.1	GI:12387774	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	1 (bases 1 to 896)		
COMMENT	NIH-MGC http://mgi.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLM10103 row: h column: 21 High quality sequence stop: 652. Location/Qualifiers		
FEATURES			
source	1..896	/organism="Homo sapiens"	


```

RESULT 11
BG115319      831 bp      mRNA      linear      EST 30-JAN-2001
LOCUS        602316274P1 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:4416354 5',
DEFINITION   mRNA sequence.
ACCESSION    BG115319
VERSION      BG115319.1 GI:12608825
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10147 row: d column: 19
High quality sequence stop: 731.
FEATURES     Location/Qualifiers
             1..831
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:4416354"
             /tissue_type="duodenal adenocarcinoma, cell line"
             /lab_host="DH10B (phage-resistant)"
             /clone_lib="NIH_MGC_88"
             /note="Organ: small intestine; Vector: pCMV-SPORT6;
             Site 1: NotI; Site 2: SalI; Cloned unidirectionally;
             oligo-dT primed. Average insert size 1.767 kb. Library
             enriched for full-length clones and constructed by Life
             Technologies. Note: this is a NIH_MGC Library."
ORIGIN
Query Match      34.5%; Score 671.8; DB 12; Length 831;
Best Local Similarity 97.4%; Pred. No. 5.7e-159;
Matches 713; Conservative 0; Mismatches 15; Indels 4; Gaps 3;

Qy 1058 CTCCTTTATGACGGCTCCGACGAGAGATGGTGCAGGTGTTATCCCGCCGAGGCGAGT 1117
Db 1 CTCCTTTATGACGGCTCCGACGAGAGATGGTGCAGGTGTTATCCCGCCGAGGCGAGT 60

Qy 1118 GGGCGCCATCATCGGCAGAGAGGGGCGAGCAGATCAACAGCTCTCCCGGTTTGGCCAGCGC 1177
Db 61 GGGCGCCATCATCGGCAGAGAGGGGCGAGCAGATCAACAGCTCTCCCGGTTTGGCCAGCGC 120

Qy 1178 CTCCTCAAGATTGCACACCGGAAACCTGACTCCAAAGTTCGTATGTTATCATCATC 1237
Db 121 CTCCTCAAGATTGCACACCGGAAACCTGACTCCAAAGTTCGTATGTTATCATCATC 180

Qy 1238 TGGACCGCCAGAGGCCCAATTCAAGCTCAGGAGAAATCTATGGCAAACTCAAGAGGA 1297
Db 181 TGGACCGCCAGAGGCCCAATTCAAGCTCAGGAGAAATCTATGGCAAACTCAAGAGGA 240

Qy 1298 GAACCTCTTTGTTCCCAAGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATC 1357
Db 241 GAACCTCTTTGTTCCCAAGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATC 300

Qy 1358 AGCAGCTGCCCGGTCATTGGCAAGGTTGGAAGAAACCGTGAACAGTTGCAAGATTGAC 1417
Db 301 AGCAGCTGCCCGGTCATTGGCAAGGTTGGAAGAAACCGTGAACAGTTGCAAGATTGAC 360

Qy 1418 GGCAGCTGAGGTGGTGTAGTACCAAGACAGACCCCTGATGAGAACGACCGAGTCACTCGT 1477

```

```

Db 361 GGCAGCTGAGGTGGTGTAGTACCAAGAGACCCAGACCCCTGATGAGAACGACCGAGTCACTGT 420

Qy 1478 GAAATCATCGGACATTTCTATGCGAGTCAGATCGCTCAACGAGAGATCCGAGACATCCT 1537
Db 421 GAAATCATCGGACATTTCTATGCGAGTCAGATCGCTCAACGAGAGATCCGAGACATCCT 480

Qy 1538 GGGCCAGGTTAAGCAGCAGCATCAGAGGGGACAGAGTAACAGGCGCCAGGACGAGGAA 1597
Db 481 GGGCCAGGTTAAGCAGCAGCATCAGAGGGGACAGAGTAACAGG-CCAGGACGAGGAA 539

Qy 1598 GTGACCGAGCCCTCCCTGTCCTTTNGAGTCCAGGACAAACACGCGGACAGAAATCGAGAGTG 1657
Db 540 GTGACCGAGCCCTCCCTGTCCTTTGAGTCCAGGACAAACACGCGGACAGAAATCGAGAGTG 599

Qy 1658 TGCTCTCCCG-GGCAGGCTCAGATGAGTGGGATCGGACACNTGGCGCGGCTGTGA 1716
Db 600 TGCTCTCCCGGCGAGGCTCAGATGAGTGGGATCGGACACNTGGCGCGGCTGTGA 659

Qy 1717 GATCAGGTTTCCCTGATTTGAGAAAGATGTTCCAGTGAGGAAACCTGATCTNTCAGC 1776
Db 660 GATCAGGTTTCCCT-ACCTGATGAGAAAGATGTTCCAGTGAGGAAACCTGATCTCTAGCC 717

Qy 1777 CCCAAACACCCA 1788
Db 718 CAAAACACCA 729

RESULT 12
BX401619/c
LOCUS        BX401619 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION   Homo sapiens cDNA clone CSODL005Y119 3-PRIME, mRNA sequence.
ACCESSION    BX401619
VERSION      BX401619.1 GI:30631999
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8942.f For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSODL005CF10NP1&cluster=8942.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CSODL005CF10NP1.
FEATURES     Location/Qualifiers
             1..1201
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="CSODL005Y119"
             /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
             /cell_line="RAMOS CELL LINE"
             /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
             25-NORMALIZED"
             /note="1st strand cDNA was primed with a NotI-oligo(dT)
             primer. Five prime end enriched, double-strand cDNA was
             digested with Not I and cloned into the Not I and EcoR V
             sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match      34.5%; Score 670.4; DB 13; Length 1201;
Best Local Similarity 84.6%; Pred. No. 1.6e-158;

```

```
QY 1051 CCTATAGCTCTTTATGAGGCTCCGAGCAGGAGATGGTGACAGTGTATATCCCGCCC 1110
Db 1448 CTCCTTACCACAAATTTGAGCAATCAGAGACGGAGACTGTGCACTCTTTATCCCGCCC 1507
QY 1111 AGCAGTGGGGCCCATCATCGGCAAGAGGAGGAGCAGCATCAACAGCTCTCCCGTTTG 1170
Db 1508 TGTCCGTTGGCGCATCATTTGGCAAGCAGGCGCCACACATCAACACAGCTTCTCGCTTG 1567
QY 1171 CCAGCGCTTCATCAAGATTGACACCCCGAAGACACCTGACTCCAAAGTTCGTATGTTA 1230
Db 1568 CGGAGCTTCGATTAAAGATCGCTCCAGCAGAGACCCAGATGCTAAAGTGGGATGTA 1627
QY 1231 TCATCACTGGACGCCAGAGGCCCAATTCAGGCTCAGGAGAGAAATCTATGCAAACTCA 1290
Db 1628 TTATCACTGGACGCCAGAGGCTCAGTTCAGGCTCAGGAGAGAAATTTATGGAATAA 1687
QY 1291 AGGAGGAACTCTTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATAGCTGTC 1350
Db 1688 AAGAAGAAACTTTGTTAGTCTTAAGAGAGAGGTGAATCTTGAAGTCAATCAGATGC 1747
QY 1351 CAGCATCAGAGCTGGCGGCTCATTTGGCAAGGTGAAAGACGGTGAACGAGTTGCAGA 1410
Db 1748 CGTCTTGTCTGCTGGCAGAGTATTGGGAAGAGGAGCAAAACGGTGAATGAGTCCAGA 1807
QY 1411 ATTGACGGCAGCTGAGTGTGTAGTACCAAGAGACAGACCCCTGATGAGAAACAGCAGG 1470
Db 1808 GTTATCAAGTGTGAAGTGTGTGTCGCCGTGACCCAGACACCTGATGAGAAATGATCAAG 1867
QY 1471 TCATCGTGAATAATCATCGGACATTTCTATCCAGTCAGATGGCTCAACGGAAGATCCGAG 1530
Db 1868 TAGTGTCAAAATAAATGCGCACTTCTATCTTTGCCAGGTTGCCAGAGAAATTCAGG 1927
QY 1531 ACATCTGGCGGAGGTTAAGCAGCAGCATCAGAG 1565
Db 1928 AAATCTGACTAGGTAAGCAGCAGCAGCAGCAG 1962

RESULT 10
BO647360
LOCUS
DEFINITION
AGENCOURT_8443628 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6284231
5', mRNA sequence.
ACCESSION
BO647360
VERSION
BO647360.1 GI:21771532
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: CGAP (Stanford)
cDNA Library Preparation: Ruben Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLC2481 row: h column: 24
High quality sequence stop: 651.
Location/Qualifiers
1..953
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6284231"
/tissue_type="hepatocellular carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"

FEATURES
source
```

```
/clone.lib="NIH_MGC_100"
/notes="Organ: liver; Vector: pOT7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
```

ORIGIN

```
Query Match 34.9%; Score 678.6; DB 13; Length 953;
Best Local Similarity 95.1%; Pred. No. 1.1e-160;
Matches 765; Conservative 0; Mismatches 30; Indels 9; Gaps 6;

QY 108 AGGAGCGGGAATTCAAATCCGAAATATTCACCCCGAGCTCCGATGGGAGTACTGGAC 167
Db 138 AGGAGCGGGAATTCAAATCCGAAATATTCACCCCGAGCTCCGATGGGAGTACTGGAC 197
QY 168 AGCTGCTGGCTCAGTATGTAAGTACAGTAGAGAACTGTGAGCAAGTGAACACCGAGAGTGAG 227
Db 198 AGCTGCTGGCTCAGTATGTAAGTACAGTAGAGAACTGTGAGCAAGTGAACACCGAGAGTGAG 257
QY 228 ACGGCAAGTGGTGAATGTCACTTATTCACCCGGGAGCAGACCCAGGCAAGCCATCATGAAG 287
Db 258 ACGGCAAGTGGTGAATGTCACTTATTCACCCGGGAGCAGACCCAGGCAAGCCATCATGAAG 317
QY 288 CTGAATGGCCACCCAGTTGGAGAACCATGCCCTCAAGGTCTCTTACATATCCCGATGAGCAG 347
Db 318 CTGAATGGCCACCCAGTTGGAGAACCATGCCCTCAAGGTCTCTTACATATCCCGATGAGCAG 377
QY 348 ATAGCACAGGACCTGAGAAATGGCGCCGAGGGGGCTTTGGCTCTCGGGGTCAAGCCCGCG 407
Db 378 ATAGCACAGGACCTGAGAAATGGCGCCGAGGGGGCTTTGGCTCTCGGGGTCAAGCCCGCG 437
QY 408 CAGGGCTCACCTGTGGCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 467
Db 438 CAGGGCTCACCTGTGGCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 497
QY 468 CGGCTCCTGGTGGCCACCCAGTATGGGTGCCATTATTTGGCAAGAGGGGGGGGGGGGGGGGG 527
Db 498 CGGCTCCTGGTGGCCACCCAGTATGGGTGCCATTATTTGGCAAGAGGGGGGGGGGGGGGGGG 557
QY 528 CGCAACATCAAAAACAGACCCAGTCCAAAGATAGACGTGATAGGAGGAGAGAACGAGGT 587
Db 558 CGCAACATCAAAAACAGACCCAGTCCAAAGATAGACGTGATAGGAGGAGAGAACGAGGT 617
QY 588 GCAGCTGAAAAGCCATCAGTGTGACCTCCAGCCCTGAGGGGTGCTCTCCGCTTGTAAAG 647
Db 618 GCAGCTGAAAAGCCATCAGTGTGACCTCCAGCCCTGAGGGGTGCTCTCCGCTTGTAAAG 677
QY 648 ATGATCTTGGAGATTATGATAAAGAGGTAAGGACACCAAAA-CGGCTGACGAGGTTC 706
Db 678 ATGATCTTGGAGATTATGATAAAGAGGTAAGGACACCAAAAACCGCTGACGAGTTC 737
QY 707 CTGTAAGATCTTGGCCCATATAAAC-TTGTAGGGGTCTCATTTGGCAAGAGGAGACGGA 765
Db 738 CTGTAAGATCTTGGCCCATATAAAC-TTGTAGGGGTCTCATTTGGCAAGAGGAGACGGA 797
QY 766 A-CCTGAGAGAGGTAGCAAGATA-CCGAGACAAAATCACCATCTCTCGTT-GCAAG 822
Db 798 ACCCTGAAGAGGTANAGCCAGATACCCGAGACAAAATACCATCTCTCTCGTTGGCAGA 857
QY 823 ACCTTACCTTTACAAACCTG---AGAGGACCATCTGTGAGGGGGGGGGGGGGGGGGGGGG 878
Db 858 ACCTTACCTTTACCAACCTTGGAGAGGAGACATCCACTGTGAAGGGGGGGGGGGGGGGGG 917
QY 879 TGTTCAGGGCGGAGCAGGAATA 902
Db 918 TGTTCAGGGCGGAGCAGGAATA 941
```



```

Db      1846 TAGTGTCAAATAAATGCGCACTTCTATGTTGCCAGGTTGCCAGAGGAATTCAGG 1905
QY      1531 ACATCTGGCCCGAGGTTAAGCAGCAGATCAGAAG 1565
Db      1906 AAATTCTGACTCAGGTAAGTAAGCAGCAGCAGCAG 1940

RESULT 8
LOCUS   BX327672/c
DEFINITION
BX327672 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
cDNA clone CS0BC018Y118 5-PRIME, mRNA sequence.
ACCESSION
VERSION   BX327672.1 GI:30342696
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 919)
  Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
  Full-length cDNA libraries and normalization
  Unpublished (2001)
  Contact: Genoscope
  Genoscope - Centre National de Sequencage
  BP 191 91006 EVRY cedex - France
  Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
  Library was constructed by life Technologies, a division of
  Invitrogen. This sequence belongs to sequence cluster 8684.r For
  more information about this cluster, see
  http://www.genoscope.cns.fr/
  cgi-bin/cluster.cgi?seq=CS0BAG007ZH09_CS00648_1&cluster=8684.r.
  Contact : Feng Liang Email : fliang@lifetech.com URL :
  http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
  Faraday Avenue Genoscope sequence ID : CS0BAG007ZH09_CS00648_1.
FEATURES
    Source
      1. 919
        /location=Qualifiers
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="CS0BC018Y118"
        /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
        /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
        /notes="1st strand cDNA was primed with a NotI-oligo(dT)
        primer. Five prime end enriched, double-strand cDNA was
        digested with Not I and cloned into the Not I and EcoR V
        sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match      37.1%; Score 722.6; DB 13; Length 919;
Best Local Similarity 98.4%; Pred. No. 7.8e-172;
Matches 750; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY      108 AGAGCGGAAATTCAAATCCG-AAATATTCACCCCGAGTCCGATGGAAAGTACTGGA 166
Db      783 AGGACCGGAAATTCAAATCCGAAATATTCACCCCGAGTCCGATGGAAAGTACTGGA 724
QY      167 CAGCTCTGGCTCAGTATGGTACAGTAGAAGTCTGAGCAAGTGAACACCGAGTGGA 226
Db      723 CAGCTCTGGCTCAGTATGGTACAGTAGAAGTCTGAGCAAGTGAACACCGAGTGGA 664
QY      227 GACGGCAGTGGTGAATGTCACTTATTCACCGGGAGCAGACCGAGCAAGCCATCATGAA 286
Db      663 GACGGCAGTGGTGAATGTCACTTATTCACCGGGAGCAGACCGAGCAAGCCATCATGAA 604
QY      287 GCTGAATGCCACCAAGTTGGAGAACCATGCCCTGAAAGGTTCTCTCATATCCCGATGAGA 346
Db      603 GCTGAATGCCACCAAGTTGGAGAACCATGCCCTGAAAGGTTCTCTCATATCCCGATGAGA 544
QY      347 GATAGCAGGAGCCTGAGAAATGGGCGCCGAGGGGCTTTGGCTCTCGGGGTACGCCCG 406
Db      543 GATAGCAGGAGCCTGAGAAATGGGCGCCGAGGGGCTTTGGCTCTCGGGGTACGCCCG 484

```

```

QY      407 CAGGGCTACCTGTGGCAGCGGGGGCCCGAGCCAAAGCAGCAGAAAGTGGACATTCGCCCT 466
Db      483 CAGGGCTCACCTGTGGCAGCGGGGGCCCGAGCCAAAGCAGCAGAAAGTGGACATTCGCCCT 424
QY      467 TCGGCTCTCTGGTCCCGCCACCGATATCTGGTGCATTATTATTTGGCAAGGAGGGGCCACCAT 526
Db      423 TCGGCTCTCTGGTCCCGCCACCGATATCTGGTGCATTATTATTTGGCAAGGAGGGGCCACCAT 364
QY      527 CGCAACATCATCAAAAACAGACACCGAGTCCCAAGATAGACGTGCATAGGAAGGAGACGAGG 586
Db      363 CGCAACATCATCAAAAACAGACACCGAGTCCCAAGATAGACGTGCATAGGAAGGAGACGAGG 304
QY      587 TGCAGCTGAAAAGCCCATCAGTGTGCATCTCCACCCCTGAGGGCTGCTCTCCGCTTGTA 646
Db      303 TGCAGCTGAAAAGCCCATCAGTGTGCATCTCCACCCCTGAGGGCTGCTCTCCGCTTGTA 244
QY      647 GATGATCTTTGGAGATTATGCATAAAGAGGCTAAGGACACCAAAACCGCTGACGAGTTCC 706
Db      243 GATGATCTTTGGAGATTATGCATAAAGAGGCTAAGGACACCAAAACCGCTGACGAGTTCC 184
QY      707 CTGGAAGATCTCGGCCCATTAATACTTTGTAGGGCTCTCATTTGGCAAGGAGACGAGAA 766
Db      183 CTGGAAGATCTCGGCCCATTAATACTTTGTAGGGCTCTCATTTGGCAAGGAGACGAGAA 124
QY      767 CTGGAAGAGGTAGACAGATACCGAGACCAAAATCACAATCTCTCTGTTGCAAGACCT 826
Db      123 CTGGAAGAGGTAGACAGATACCGAGACCAAAATCACAATCTCTCTGTTGCAAGACCT 64
QY      827 TACCCCTTTACAAACCTTGAGAGGACCATCACTGTGAAGGGGGC 868
Db      63 TACCCCTTTACAA-CTTGAGAGGACCATCACTTTGAAGGGGGC 23

RESULT 9
LOCUS   AK011689
DEFINITION
Mus musculus 10 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:2610036H8 product:insulin-like growth
factor 2, binding protein 3, full insert sequence.
ACCESSION
VERSION   AK011689.1 GI:12847972
KEYWORDS   HTC; CAP trapper.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
  1
  Carninci, P. and Hayashizaki, Y.
  High-efficiency full-length cDNA cloning
  Meth. Enzymol. 303, 19-44 (1999)
  99279253
  10349636
REFERENCE
  2
  Carninci, P. and Hayashizaki, Y.
  Normalization and subtraction of cap-trapper-selected cDNAs to
  Prepare full-length cDNA libraries for rapid discovery of new genes
  Genome Res. 10 (10), 1617-1630 (2000)
  20499374
  11042159
REFERENCE
  3
  Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
  Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
  Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
  Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
  Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
  Yoneda, Y., Iehikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
  Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
  RIKEN integrated sequence analysis (RISA) system--384-format
  sequencing pipeline with 384 multipillar sequencer
  Genome Res. 10 (11), 1757-1771 (2000)
  20530913
  11076861

```

Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
Trust/MRC building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>.

FEATURES

Location/Qualifiers
1. 4100
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="PANTOM,DB:E430018F06"
/db_xref="MGI:247362"
/db_xref="taxon:10090"
/clone="E430018F06"
/cell_type="thymic cells"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days neonate"
243..1982
/note="unnamed protein product; insulin-like growth factor
2, binding protein 3 (MGI:MGI:1890359, GB|NM_023670,
evidence: BLASTN, 99%, match=1740)
putative"
/codon_start=1
/protein_id="BAC40370.1"
/db_xref="GI:26353480"
/translation="MNKLYIGNLSHDHAGPADLESVFKDAKIPVAGPFLVKTGYAFVDC
PDEGWLKALSKMELHGKMEVHSKPRQIRKQIRNPPLHQLWEVLDSLLV
QYGVSECOVNTSEAVNVTVSSKQARQALDKLNGFOLENTLKVAIPIPETAA
QONSPQLRGRGQRGSSQASPSVSKQPCDLRLILVPTQFVCAIIGKEGATI
RNIKTQGLDKVHKNKIGKEINLKKIEQDTDKITISPLQELTINPRTILVKSVE
IPLKILAHNNFVGLIGKEINLKKIEQDTDKITISPLQELTINPRTILVKSVE
LTPKAEIIMKIRESENDIASMNLQANLIPGNLNLGFLPPTSGMPPTSPST
LTAPYPOQESQETVHLFIPALSGAIIGQGHILKOLSRFAGSIKIAABAPDAP
VRMIIITGPPPAQFAQGRIVYKIKENFVSPKEVLEAHIRVPSFAGRVIGKGGK
TVNLSLSLSVAVVPRDQTPDENQVVKITGHFYACQVAQRKIQEILITQVKHQHQ
KALQSGFPQSRKK"
4079..4084
/note="putative"
4100
/note="putative"

CDS

243..1982
/note="unnamed protein product; insulin-like growth factor
2, binding protein 3 (MGI:MGI:1890359, GB|NM_023670,
evidence: BLASTN, 99%, match=1740)
putative"

polyA_signal

4079..4084
/note="putative"

polyA_site

4100
/note="putative"

ORIGIN

Query Match 37.2%; Score 723.4; DB 11; Length 4100;
Best Local Similarity 68.7%; Pred. No. 1.1e-171;
Matches 1013; Conservative 0; Mismatches 456; Indels 6; Gaps 1;
QY 97 CTTCCCGGTTAGGAGCGGAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGG 156
DB |||||
QY 466 CTAACGGCAGAGGATTCGTAACTTCAGATACGAAATATCCGCCCACTTACATGGG 525
DB |||||
QY 157 AAGTACTGGACAGCGCTGCTGCTCAGTATGGTACAGTAGAGAACTGTGAGCAAGTGAACA 216
DB |||||
QY 526 AGGTCTGGATAGTTTACTAGTCCAGTATGGATGGTGGAGAGCTGTGAGCAAGTGAACA 585
DB |||||
QY 217 CCGAGAGTGGACGCGAGTGGTGAATGTCACTATTCCACCGGAGCGAGCACGAGCAAG 276
DB |||||
QY 586 CGGATTCGAAACGGCAGTGTGAATGTAAATGTAACTATTCCAGTAGAGGACCAAGCTAGACAG 645
DB |||||
QY 277 CCATCATGAGCTGAATGGCCACAGTTGGAGAACCATCGCTCGTGAAGTCTCTATCATCC 336
DB |||||
QY 646 CTTTACACAACTGAATGATTCAGTTAGAGAACTTACCTTGAAGTTGCTTACATCC 705
DB |||||
QY 337 CCGATGACGATAGCAGGAGCACTGAGAAATGGGCGCGAGGGGCTTTTGGCTCTCGGG 396
DB |||||
QY 706 CAGATGAAGTCTGCTGCCAGCAAAATCCCTCCACACAGCTCCCGGGGCGCGGGGCCAG 765
DB |||||
QY 397 GTACGCCCGCAGGGCTCACCTGTGGCAGCGGGGGCCCCAGCCAGCAGCAGCAGTAGTG 456

Db 766 GGCAGCGGGGCTCATCCAGGCGGCTCTCCAGGATCGGTGTCAAGCAGAAACCTCTGT 825
QY |||||
QY 457 ACATCCCCCTTTCGGCTCTCTGTGTCGCCACCCAGTAGTATGGGTGCCATTTATGGCAGGAGG 516
Db |||||
Db 826 ACCTGCCACTGCGGCTCTGTGTTCCACCCAGTGTGTTGGAGCCATTTATAGGAAAGAG 885
QY |||||
QY 517 GGGCCACATCCGCAACATCACAACACAGACCCAGTCCAGATAGACCTGATAGGAGG 576
Db |||||
Db 886 GTGCTACCATTCGCAACATCACAAGCAGACCCAGTCTAAATTCATGATCGATCGTAAAG 945
QY |||||
QY 577 AGACGCGAGGTGACGCTGMAAAGCATCAGTGTGCACTCCACCCCTGAGGCGCTCTCT 636
Db |||||
Db 946 AGAATACAGGGCGCGGAGAGTCCATTACTCTCTCTACCCCTGAGGCACCTCTG 1005
QY |||||
QY 637 CCGCTTTGTAAGATCATCTTGAGATATTCATTAAGAGGCTTAAGGACACCAAAACGGCTG 696
Db |||||
Db 1006 CAGCTTGTAGTCTATTCTGGAGATTATGCATAAGGAAGCTCAAGATATAAATTCACAG 1065
QY |||||
QY 697 ACGAGTTCCTCCCTGAGATCTGCGCCATATTAATTTGTAGGCGCTCTCATTTGGCAAGG 756
Db |||||
Db 1066 AGGAGATTCCCTTGAAGATCTTAGCTCAATAACTTTGTAGGCGCTCTCATTTGTAAG 1125
QY |||||
QY 757 AAGCAGGAACTCTGAAGAGGTAGAGCAAGATACCGAGACAAAATCAACATCTCTCT 816
Db |||||
Db 1126 AAGGAGAACTTTAAATAATCGAGCAAGACACGACACTAAATCAATATCTCCAT 1185
QY |||||
QY 817 TGCAGACCTTACCTTTACAACTTGAGAGACCATCATCTGTGAAGGGGCCATCGAGA 876
Db |||||
Db 1186 TGCAGGAATCTGACGTGTACAACTCGGAAACGACCATTTACAGTGAAGAGCAGTGTGAGA 1245
QY |||||
QY 877 ATTGTTGAGGGCGGAGCAGCAAAATTAAGAAAGTTCGGGAGGCGCTTATCAGAAATGATG 936
Db |||||
Db 1246 CGTGTGCCAAGCGGAGGAGAAATTAATGAAGAGATCAGGAGGCTTTATGAAATGATA 1305
QY |||||
QY 937 TGGCTGCCATGAGCTCTC-----ACCTGATCCCTGGCGCTGAAACCTGGCTGTGTAGGTC 990
Db |||||
Db 1306 TTGCTTCATGAACTTCAAGCAATTTAAATCCCTGGATTAATCTGATGCTTGGGTC 1365
QY |||||
QY 991 TTTTCCAGCTTCATCCAGCGAGCTCCGCGCGCTCCCGAGAGCGTTCAGGGGCTGCTC 1050
Db |||||
Db 1366 TGTTCACCCAGCTCAGGATGCCACCTCCACCTCAGGCGCCCTTCAACCCCTGACTC 1425
QY |||||
QY 1051 CCTATAGCTCTTTATGAGGCTCCCGAGCAGGAGATGTCAGTGTGTTATCCCGGCC 1110
Db |||||
Db 1426 CTCCTACCCCAATTTGAGCAATCAGAGACGAGACTGTGATCTGTGTTATTCGCGCC 1485
QY |||||
QY 1111 AGCAGTGGGGCCCATCATCGGCAAGAGGGGCGACACATCAAAACAGCTCTCCCGTTTG 1170
Db |||||
Db 1486 TGTCCGTTGGGCCCATCATTCGCAAGCAGGGCCACACATCAAAACAGCTTCTCGCTTG 1545
QY |||||
QY 1171 CCAGCGCTTCATCAAGATTGACACCCCGAAACACCTGACCTCCAAAGTTCGTATGGTTA 1230
Db |||||
Db 1546 CGGAGCTTCGATTAAAGATCGCTCCAGCGCAAGCACCAGATGCTAAAGTCGGGATGGTGA 1605
QY |||||
QY 1231 TCATCACTGGACCGCAGGCGCCAAATTCAGGCTCAGGAGAGATCTATGCAAACTCA 1290
Db |||||
Db 1606 TTATCACTGGACCAACAGAGGCTCAGTTCAGGCTCAGGAGAGAAATTTATGAAATAATTA 1665
QY |||||
QY 1291 AGGAGGAGAACTTCTTTGCTCCAAAGGAGAAAGTGAAGCTGGAGAGCCACATACGTTGC 1350
Db |||||
Db 1666 AAGAGAGAAACTTTGTTAGTCTTAAAGAGAGTGAATCTTGAAGCTCACATCAGAGTGC 1725
QY |||||
QY 1351 CAGCATCAGAGCTGGCGGCTTATGGCAAGGTTGAAAGGAGGAGGAGGAGGAGGAGGAGGAG 1410
Db |||||
Db 1726 CGTCTCTTCTGCTGGCAGAGTTATTGGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1785
QY |||||
QY 1411 ATTTGAGCGCAGCTCAGTGTGTAGTACCAAGAGACAGACCCCTGATGAGAGACGACACAG 1470
Db |||||
Db 1786 GTTTATCAAGTCTGAAGTGTGTCGCCCGCTGACCAAGCACCTGATGAGATGATCAAG 1845
QY |||||
QY 1471 TCATCGTGAATAATCATCGACATTTCTATGCGAGTCAAGTGGCTCAACGAGAGATCCGAG 1530

ORIGIN		Library."	
Query Match		38.3%; Score 745.4; DB 13; Length 950;	
Best local similarity		94.2%; Pred. No. 1.3e-177;	
Matches		864; Conservative 0; Mismatches 36; Indels 17; Gaps 8;	
QY	264	CACACAGGCAAGCCATCATGAGCTGAATGGCCACCAAGTTGGAGAACCAATGCCCTGAAG	323
Db	1	CAGACAGGCAAGCCATCATGAGCTGAATGGCCACCAAGTTGGAGAACCAATGCCCTGAAG	60
QY	324	GTCTCTACATCCCTCCGATGACGATAGACAGGACCTGAGAAATGGCGCCGAGGGGC	383
Db	61	GTCTCTACATCCCTCCGATGACGATAGACAGGACCTGAGAAATGGCGCCGAGGGGC	120
QY	384	TTTGGCTCTCGGGTTCAGCCCGCCAGGGTCTACCTGTGGCAGCGGGGGCCCCAGCAAG	443
Db	121	TTTGGCTCTCGGGTTCAGCCCGCCAGGGTCTACCTGTGGCAGCGGGGGCCCCAGCAAG	180
QY	444	CAGCAGCAAGTGCATCCCTTCGGCTCTCGTGGCCACCCAGTATGTGGTGCATTT	503
Db	181	CAGCAGCAAGTGCATCCCTTCGGCTCTCGTGGCCACCCAGTATGTGGTGCATTT	240
QY	504	ATTGSCAAGAGGGGGCCACCATCCGCAACATCACAAACAGACCCAGTCCAAAGATAGAC	563
Db	241	ATTGSCAAGAGGGGGCCACCATCCGCAACATCACAAACAGACCCAGTCCAAAGATAGAC	300
QY	564	GTGCATAGAGAGAGAACGAGTGCAGTGTGAAAGGCCATCATGTGTGCACTCCACCCCT	623
Db	301	GTGCATAGAGAGAGAACGAGTGCAGTGTGAAAGGCCATCATGTGTGCACTCCACCCCT	360
QY	624	GAGGGCTGTCTCCCTGCTGATAGATGATCTTGGAGATTATGATTAAGAGGCTTAGGAC	683
Db	361	GAGGGCTGTCTCCCTGCTGATAGATGATCTTGGAGATTATGATTAAGAGGCTTAGGAC	420
QY	684	ACCAAAACGGCTGACGAGGTTCCCTTGAAGATCCTGGCCCATTAATTAATTTGTAGGGCGT	743
Db	421	ACCAAAACGGCTGACGAGGTTCCCTTGAAGATCCTGGCCCATTAATTAATTTGTAGGGCGT	480
QY	744	CTCATTTGCAAGAGAGAGAGACCTTGAAGAGGTTAGACAGATACCGAGACAAAATC	803
Db	481	CTCATTTGCAAGAGAGAGAGACCTTGAAGAGGTTAGACAGATACCGAGACAAAATC	540
QY	804	ACCATCTCTCTGTGCAAGACCTTACCTTTACACCTGAGAGGACCATCT-CTGAA	862
Db	541	ACCATCTCTCTGTGCAAGACCTTACCTTTACACCTGAGAGGACCATCTCTGTGAA	600
QY	863	GGGGGCCATCGAGAAATTGTTGAGGGCCGAGCAGGAAATAATGAAGAAAGTTCCGGAGGC	922
Db	601	GGGGGCCATCGAGAAATTGTTGAGGGCCGAGCAGGAAATAATGAAGAAAGTTCCGGAGGC	559
QY	923	CTATGAGATGATGGCTGGCATGAGC-----TCTCACCTGATCCCTGGCTGAACT	976
Db	660	CTATGAGATGATGGCTGGCATGAGC-----TCTCACCTGATCCCTGGCTGAACT	719
QY	977	-GGCTGCTGTAGTCTTTTCCAGCTTCATCCAGCGCTCCCGCGCTCCAGCAGC-1034	
Db	720	GGGCTGCTGTAGTCTTTTCCAGCTTCATCCAGCGCTCCCGCGCTCCAGCAGC-779	
QY	1035	GTTACTGGGGCTGCT-CCCTATAGTCTCTTTTATGAGGCTCCCGAGCAGGATGTGTG--1091	
Db	780	GTTACTGGGGCTGCTCTCTCTATAGTCTCTTTTATGAGGCTCCCGAGCAGGATGTGTG--839	
QY	1092	CAGGTGTTTATCCCGCCAGGAGTGGGGC-----CATCATGGCAAGAGGGGAGCA1147	
Db	840	AAGGGGTTTAAACCCCGCCAGGAGTGGGGC-----CATCATGGCAAGAGGGGAGCA899	
QY	1148	CATCAACAGCTCTCC1164	
Db	900	CTCAACAGCTCTCC916	
RESULT 7			

AK088465

LOCUS

DEFINITION

AK088465

Mus musculus 2 days neonate thymic cells cDNA, RIKEN full-length enriched library, clone:E430018F06 product:insulin-like growth factor 2, binding protein 3, full insert sequence.

AK088465.1

GI:26353479

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohtsuka, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

4

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium

Functional annotation of a full-length mouse cDNA collection

Nature 409, 685-690 (2001)

5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 4100)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, T., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koyu, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

7

cdna library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken


```
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1042)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12263 row: 1 column: 24
High quality sequence stop: 682.
FEATURES
    source
        1..1042
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5550887"
            /tissue_type="retinoblastoma"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH MGC 67"
            /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
            Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
            Average insert size 1.75 kb. Library constructed by Life
            Technologies."
ORIGIN
    Query Match 42.6%; Score 828.4; DB 12; Length 1042;
    Best Local Similarity 96.8%; Pred. No. 1.2e-198;
    Matches 920; Conservative 0; Mismatches 18; Indels 12; Gaps 7;
QY 676 CTAAGGACACCAAAACGGCTGACGAGTTCCCTGAAGATCCTGGCCCATTAATCACTTGT 735
Db 2 CTCAGGACACCAAAACGGCTGACGAGTTCCCTGAAGATCCTGGCCCATTAATCACTTGT 61
QY 736 TAGGGCGTCTCATTTGGCAGGAAGGACGGAACCTGAAGAGGTAGAGCAATACCGAGA 795
Db 62 TAGGGCGTCTCATTTGGCAGGAAGGACGGAACCTGAAGAGGTAGAGCAATACCGAGA 121
QY 796 CAAAATCACCATCTCTCGTTGCAAGACCTTACCTTTACAACTCAGAGGACCATCA 855
Db 122 CAAAATCACCATCTCTCGTTGCAAGACCTTACCTTTACAACTCAGAGGACCATCA 181
QY 856 CTGTGAAGGGGGCCATCAGAAATTTGTTCAGGGCCGAGCAGGAATATGAAGAAAGTTC 915
Db 182 CTGTGAAGGGGGCCATCAGAAATTTGTTCAGGGCCGAGCAGGAATATGAAGAAAGTTC 241
QY 916 GGGAGGCGCTATGAGAAATGATGGTGGTCATGAGC-----TCTCAGCTGATCCCTGGCC 969
Db 242 GGGAGGCGCTATGAGAAATGATGGTGGTCATGAGC-----TCTCAGCTGATCCCTGGCC 301
QY 970 TGAACCTGGCTGTAGTCTTTTCCAGCTTTCATCCAGCGCAGTCCCGCGCCCTCCCA 1029
Db 302 TGAACCTGGCTGTAGTCTTTTCCAGCTTTCATCCAGCGCAGTCCCGCGCCCTCCCA 361
QY 1030 GCAGGTTTACTGGGGCTCTCCCTATAGTCTCTTTATGCAGGCTCCCGAGCAGGAGATGG 1089
Db 362 GCAGGTTTACTGGGGCTCTCCCTATAGTCTCTTTATGCAGGCTCCCGAGCAGGAGATGG 421
QY 1090 TGCAGGTTTATCCCGCCGAGCAGTGGGGCCATCATCGGCAAGAGGGGCGAGCACA 1149
Db 422 TGCAGGTTTATCCCGCCGAGCAGTGGGGCCATCATCGGCAAGAGGGGCGAGCACA 481
QY 1150 TCAACAGCTCTCCCGGTTTGCAGCGCTCCATCAAGATTGCACACCCGAAAACACCTG 1209
Db 482 TCAACAGCTCTCCCGGTTTGCAGCGCTCCATCAAGATTGCACACCCGAAAACACCTG 541
QY 1210 ACTCCAAAGTTCTGATGTTTATCATCTGACCGCCAGAGGCCCAATTTCAAGGCTCAGG 1269
Db 542 ACTCCAAAGTTCTGATGTTTATCATCTGACCGCCAGAGGCCCAATTTCAAGGCTCAGG 601
QY 1270 GAAGAAATCTATGTCACAACTCAAGGAGGAGAACTCTTTTGGTCCCAAGAGGAAGTGAAGC 1329
Db 602 GAAGAAATCTATGTCACAACTCAAGGAGGAGAACTCTTTTGGTCCCAAGGAGGAAGTGAAGC 661
QY 1330 TGGAGACCCACATACGTTGTCAGCATCAGCAGTGGCCGGGTCATTTGGCAAAAGTGGAA 1389
Db 662 TGGAGACCCACATACGTTGTCAGCATCAGCAGTGGCCGGGTCATTTGGCAAAAGTGGAA 721
QY 1390 AAACGGTGAACGAGTTCAGAAATTTGACGCGCAGCTGAGGTGGTGTAGTACCAAGAGACCA 1449
Db 722 AAACGGTGAACGAGTTCAGAAATTTGACGCGCAGCTGAGGTGGTGTAGTACCAAGAGACCA 781
QY 1450 CCCCTGATGAGAACGACCATGTCATCGTGAATAATCATC-GGACATTTCTATGCCAGTCAG 1508
Db 782 CCCCTGATGAGAACGACCATGTCATCGTGAATAATCATCGGACATTTCTATGCCAGTCAG 841
QY 1509 ATGGCTCAAC-GGAAGATCCGAGACATCCT-GGCCAGGTTAAGCAGCAGCAT-CAGAG 1565
Db 842 ATGGCTCCACGGGAAGATCCGAGACATCCTGGGGCCGAGGTTAAGCAGCAGCATCCAAAAG 901
QY 1566 GGACAGA-GTAACCCAGGCCAGGCAC-GGAGGAAGTGACCAAGCCCTCC 1613
Db 902 GGACAGAGGTTAACCCAGGCCAGGCACGGGAGGAAAGTGAACAGCAGCATCCAAAAG 951
RESULT 6
LOCUS BQ648383 950 bp mRNA linear EST 15-JUL-2002
DEFINITION AGENCOURT 8342265 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6268202
5', mRNA sequence.
ACCESSION BQ648383
VERSION BQ648383.1 GI:21772555
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 950)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CGAP (Stanford)
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2439 row: m column: 03
High quality sequence stop: 599.
FEATURES
    source
        1..950
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:6268202"
            /tissue_type="neurocellular carcinoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH MGC 100"
            /note="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2:
            EcoRI; cDNA made by oligo-dT priming. Directionally cloned
            into EcoRI/XhoI sites using the following 5' adaptor:
            GGCACGAG(G). Size-selected >500bp for average insert size
            1.8kb. Library constructed by Ling Hong in the laboratory
            of Gerald M. Rubin (University of California, Berkeley)
            using ZAP-cDNA synthesis kit (Stratagene) and Superscript
            II RT (Life Technologies). Note: this is a NIH_MGC
```


URL: http://genome.gsc.riken.go.jp/ , Tel: 81-45-503-9222, Fax: 81-45-503-9216	
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAAGGATCCAGAGCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGATCTCGAGTTAATTAATCCGCCGCCGCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.	
FEATURES	Location/Qualifiers
source	1. .2780 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM:DB:3110001D23" /db_xref="MGI:1904502" /db_xref="taxon:10090" /clone="3110001D23" /tissue_type="head" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="13 days embryo" 279..2012 /note="unnamed protein product; insulin-like growth factor 2, binding protein 1 (MGD MGI:1890357) putative" /codon_start=1 /protein_id="BAB29071.1" /db_xref="GI:12851514" /translation="MNKLYIGNLNEVTPADLEKVFAEHKISYSGOFLVKSGYAFVDC PDHMAKAIETFGKVELQKRIEHSVPKQKRSKIQIRNIPQLRWEVLDSLLA QYPTNCEQVNTSESTAVVNTSYNRQETROAKMLNGHOLENHALKYSYIPDEQIT QGPENRRGGGSRGQPGSPVAAGAPAKQPDVPLRLIVPTQVUGALIEKGATIRNITKQTKLIDVHRNENAEKALSVHSTPBGSSACHMILEIMHKAEDTKTADIE VPKLIHANHFVRLKIGKRNLLKVEQDTETKITISSLDLTYNPERTIVKGAIE NCRABQEMKKRYENDVAAVMSQLSHLIPGLNLAAGVLFPAASSAVPPPPSVTVG AAPPYGFMAQEQVQVFIAPAQVAGAILGKGQHOIKQLSRFASIKIAPPETPSK VRMVITGPPEAQKQAFQIYGLKEENFEPKPEVKLETHIRVPASAGRVIGKSGK TVNELQNLTAEEVVVRDQTPDENDQVIVKLIIGHFVYASQAKIRIDILAQVQKQHK GQSNLQAARRK"
CDS	
ORIGIN	
Query Match 66.8%; Score 1299.2; DB 11; Length 2780; Best Local Similarity 83.3%; Pred. No. 0; Matches 1556; Conservative 0; Mismatches 278; Indels 33; Gaps 6;	
QY	108 AGAGCGGGAATTCGAATCCGAATATTCACCCAGCTCCGATGGGAAGTACTGGAC 167
Db	513 AGAGTGGGAAATACAGATTCCGAATATTCACCTCAGCTCCGATGGGAAGTACTGAT 572
QY	168 AGCTGTGTGCTCAGTATGTACAGTAGAGAACTGTGTGAGCAAGTGAACCCGAGAGTGAG 227
Db	573 AGCTGTGTGCTCAGTACGTACGTGGGAACCTGTGAGCAAGTGAACACTGAAAGTGAG 632
QY	228 ACGCAGTGTGTGAATGTCACTATTCCACCGGAGCAGACCCAGGCAAGCCATCATGAAG 287
Db	633 ACAGCTGTGTGCTCAACGTCACTACTTAACCGGAGCAGACCCAGGCAAGCTATCATGAAG 692
QY	288 CTGAATGGCCACAGTTGGAGAACCATGCCCTGAAGGTCTCTTACATCCCCGATGAGCAG 347
Db	693 CTAAATGGCCATCAACTGGAGAACCATGCCCTGAAGGTCTCTTACATFACCTGATGAGCAG 752
QY	348 ATAGCACAGGACCTGAGATGGCGCGGAGGGGCTTTGGCTCTCGGGTCAAGCCCGC 407
Db	753 ATACCGAAGTCTCTGAGATGGCGGTCTGTGGAGGGCTTTGGGTCTCTCGGGCCAGCCCGG 812
QY	408 CAGGGCTCACCTGTGGCAGCGGGGCCCCAGCAAGCAGCAGCAAGTGGACATCCCCCTT 467

REFERENCE AUTHORS	3 Shibata, K., Itoh, M., Aizawa, K., Nagasaka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, U., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 20530913 PUBMED 11076861
REFERENCE AUTHORS	4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)
REFERENCE AUTHORS	5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 2718)
REFERENCE AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akai, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
REFERENCE AUTHORS	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
REFERENCE AUTHORS	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Tissues were provided by Dr. Tomohiro Kono (Department of Animal Science, Tokyo University of Agriculture, 1737 Hunko Atsugi City, Kanagawa Prefecture, Japan) whose assistance we gratefully acknowledge. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/ URL: http://fantom.qualifiers
FEATURES source	location/Qualifiers 1..2718 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:B130008A04" /db_xref="MGI:2410206" /db_xref="taxon:10090" /clone="B130008A04" /issue_type="parthenogenote" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="9.5 days embryo" 217..1950 /note="unnamed protein product; insulin-like growth factor 2, binding protein 1 (MGD MGI:1890357) putative"
CDS	
ORIGIN	Query Match 66.8%; Score 1299.2; DB 11; Length 2718; Best Local Similarity 83.3%; Pred. No. 0; Matches 1556; Conservative 0; Mismatches 278; Indels 33; Gaps 6; QY 108 AGAGCGGAAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTACTGGAC 167 Db 451 AGAGTCGGAATAACAGATCCGCAATATTCACACCTCAGCTCCGATGGGAAGTCTAGAT 510 QY 168 AGCTCTGCTGCTCAGTATGTCAGTAGAGACTGTGACCACTGACCACTGACCACTGAG 227 Db 511 AGCTCTGCTGCTCAGTATGTCAGTAGAGACTGTGAGCACTGAGCACTGAAAGTGAG 570 QY 228 ACAGCAGTGTGTAATGTCACCTATTCCACCGGAGCAGACCAAGCAAGCATCATGAAG 287 Db 571 ACAGCTGTGTCACTCACTACTCTAACCGGAGCAGACCAAGCAAGCATCATGAAG 630 QY 288 CTGAATGGCCACCACTGTGGAGAACCAATGCCCTGAAGGTCTCTACATPCCCGATGACGAG 347 Db 631 CTAATGGCCATCACTGGAGAACCACTGCCCTGAAGGTCTCTCTACATACCTGATGACGAG 690 QY 348 ATAGCACAGGACCTCAGAAATGGCGCGGAGGGGGCTTTGGCTCTCGGGGTGAGGCCCGC 407 Db 691 ATACCAAGGTCTCTGAGAAATGGCGCTCTGGAGGGCTTTGGGTCTCGGGGCGAGCCCGG 750 QY 408 CAGGGCTCACTGTGGCAGCGGGGGCCCGCAGCAAGCAGCAGCAAGTGACATCCCCCTT 467 Db 751 CAAAGGTCCCGCTGCGCAGCAGGGGCTCCAGCAAGCAGCAGCAGTGGACATCCCTCTC 810 QY 468 CGCTCTCTGTCGCCACCCAGTATGTGGGTGCCATTTATTTGGCAAGAGGGGGCCACATC 527 Db 811 CGCTCTCTGTCGCTACGCAATGTATGTAGGCGCTATCATTTGGCAAGAGGGGTGCGACATC 870 QY 528 CGCAACATCACAAAACAGACCCAGTCCAAAGATAGCTCATAGGAGAGGAGAGCAGGT 587 Db 871 CGAACATCACAAAACAGACCCAGTCCAAAGATAGCTCATAGGAGAGGAGATGCGGGC 930 QY 588 GCAGCTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGGTGCTCTCCCTGCTTAAG 647 Db 931 GCTGCGGAGAGGCCCATCAGCGTGCATTCACCCCTGAAAGGTGCTCTCTCCGCGTCAAG 990 QY 648 ATGATCTTGGAGATATGATTAAGAGGCTTAGGACACCAAAACGGCTGACAGGTTCC 707 Db 991 ATGATCTTGGAGATATGATTAAGAGGAGGAGGACCAACCAACCGCAGATGAAAGTTCC 1050 QY 708 CTGAAGATCTCGCCCATTAATTAACCTTTAGGGCGCTCTCATTTGGCAAGAGGAGGAGAAC 767 Db 1051 CTGAAGATCTCGCTCATTAACCACTTCGTCGGCGGAGCTCATTTGGCAAGAGGAGGAGAAC 1110 QY 768 CTGAAGAGGTAGACCAAGATCCGAGACAAAATACCATCTCTCTGTTGCAAGACCTT 827 Db 1111 CTGAAGAGGTGGAGCAGCAGACAGAGAGCAAGATCAACCATCTCATCGCTCCAGGACCTC 1170 QY 828 ACCCTTACCAACCTCAGAGGACCATCACTGTAGAGGGGGCCATCGAGAAATTTGTCAGG 887 Db 1171 AGCTCTATAACCCCTGAGAGGACCATCACTGTGAAGGGCGCCATTTGAGAACTGTTGACG 1230 QY 888 GCCGAGCAGGAAATTAATGAAGAAAGTTCCGGAGGCGCTATGAGAAATGATGTGGCTGCCATG 947

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2004, 06:21:04 ; Search time 4863.33 Seconds
(without alignments)
11948.985 Million cell updates/sec

Title: US-09-270-437D-7
Perfect score: 1946
Sequence: 1 gctgtacggagggtggg.....atttcctcaggttttaaaa 1946

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :			
EST:*			
1:	em_estba:*		
2:	em_esthum:*		
3:	em_estin:*		
4:	em_estmu:*		
5:	em_estov:*		
6:	em_estpl:*		
7:	em_estro:*		
8:	em_hic:*		
9:	gb_est1:*		
10:	gb_est2:*		
11:	gb_hic:*		
12:	gb_est3:*		
13:	gb_est4:*		
14:	gb_est5:*		
15:	em_estfun:*		
16:	em_estom:*		
17:	em_gss_hum:*		
18:	em_gss_inv:*		
19:	em_gss_pln:*		
20:	em_gss_vrt:*		
21:	em_gss_fun:*		
22:	em_gss_mam:*		
23:	em_gss_mus:*		
24:	em_gss_pro:*		
25:	em_gss_rod:*		
26:	em_gss_phg:*		
27:	em_gss_vrl:*		
28:	gb_gss1:*		
29:	gb_gss2:*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1299.2	66.8	2718	11	AK044850	AK044850 Mus muscu
2	1299.2	66.8	2780	11	AK013940	AK013940 Mus muscu
3	969.4	49.8	1085	13	BUL90334	BUL90334 AGENCOURT
4	921.4	47.3	1042	12	EM928196	EM928196 AGENCOURT

5	828.4	42.6	1042	12	BM561057	BM561057 AGENCOURT
6	745.4	38.3	950	13	BQ648383	BQ648383 AGENCOURT
7	723.4	37.2	4100	11	AK088465	AK088465 Mus muscu
c						
8	722.6	37.1	919	13	EX327672	EX327672 BX327672
9	721.8	37.1	2202	11	AK011689	AK011689 Mus muscu
10	678.6	34.9	953	13	BQ647360	BQ647360 AGENCOURT
11	671.8	34.5	831	12	BG115319	BG115319 602316274
c						
12	670.4	34.5	1201	13	EX401619	EX401619 BX401619
13	669	34.4	712	13	BQ225582	BQ225582 AGENCOURT
14	662.2	34.0	955	13	BQ651665	BQ651665 AGENCOURT
15	662	34.0	896	10	BF984962	BF984962 602308484
c						
16	660.2	33.9	891	13	BX372273	BX372273 BX372273
17	657	33.8	764	12	BM006191	BM006191 603614013
18	636.4	32.7	983	13	BQ643920	BQ643920 AGENCOURT
19	602.2	30.9	1248	29	AY417413	AY417413 Mus muscu
20	590.4	30.3	982	13	BQ652180	BQ652180 AGENCOURT
21	588.8	30.3	929	13	BQ652586	BQ652586 AGENCOURT
22	588.4	30.2	978	12	BG115593	BG115593 602317291
23	582.2	29.9	1173	12	BM467135	BM467135 AGENCOURT
24	576	29.6	3592	11	AK049196	AK049196 Mus muscu
25	575	29.5	2964	11	AK044984	AK044984 Mus muscu
26	557.4	28.6	1321	29	AY417411	AY417411 Homo sapi
27	554.6	28.5	3262	11	AK077404	AK077404 Mus muscu
28	547.6	28.1	884	13	BU598549	BU598549 AGENCOURT
29	546	28.1	568	14	CB146278	CB146278 K-EST0201
30	542.6	27.9	1061	13	BQ647561	BQ647561 AGENCOURT
31	540	27.7	545	13	EX089485	EX089485 BX089485
32	513.6	26.4	658	13	BQ256493	BQ256493 NISC.k003
c						
33	513.4	26.4	656	10	AW303505	AW303505 xvl19h09.x
34	513.4	26.4	1321	29	AY417412	AY417412 Pan trogl
c						
35	512.4	26.3	550	9	AA978341	AA978341 oq40a04.s
36	505	26.0	697	13	BU400982	BU400982 604139860
37	498.4	25.6	643	14	CF894671	CF894671 A0138E10-
c						
38	497.2	25.5	517	10	BE550506	BE550506 7a27f01.x
39	495.4	25.5	692	13	BU460238	BU460238 603370670
40	491.2	25.2	677	14	CB841164	CB841164 M15R-1646
41	487.4	25.0	896	13	BQ647425	BQ647425 AGENCOURT
c						
42	485.4	24.9	761	9	AW003366	AW003366 wq65h07.x
43	475	24.4	714	13	BU268759	BU268759 603820078
c						
44	469	24.1	484	9	AI337147	AI337147 qx83c08.x
45	459.4	23.6	486	9	AL704124	AL704124 DKF25686B

ALIGNMENTS

RESULT 1

AK044850

LOCUS

DEFINITION

AK044850

2718 bp mRNA linear HTC 20-SEP-2003

Mus musculus 9.5 days embryo parthenogenote cDNA, RIKEN full-length

enriched library, clone:B130008A04 product:insulin-like growth

factor 2, binding protein 1, full insert sequence.

AK044850

AK044850.1 GI:26336870

VERSION

HTC; CAP trapper.

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

JOURNAL

99279253

PUBLISHED

10349636

REFERENCE

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL

20499374

MEDLINE

PUBLISHED

11042159

This Page Blank (uspto)

Db 830 TCCCTTGAAGATTCTTACATATAAATTTGTTGGACGCTTATTTGGTAAAGAGGAA 889
QY 763 GGAACCTCAAGAAGTTAGAGCAAGATACCGAGCAAAAATCACCATCTCCTGTTGCAAG 822
Db 890 GAATCTTTAAAAAATTGAGCAAGACACAGACACTTAAATCAGATATCTCCATTGCAGG 949
QY 823 ACCTTACCCCTTTACAACTCTGAGAGACCATCCTGTGAAGGGGCCATCGAGAAATTGTT 882
Db 950 AATTGACGCTGTATTAATCCAGAACGCACTATTACAGTTTAAAGGCAATGTTGAGACATGTG 1009
QY 883 GCAGGGCCGAGCAGAGAAATATGAAGAAAGTTCCGGAGGCCCTATGAGATGATGTGGCTG 942
Db 1010 CCAAAGCTGAGAGGAGATCATGAAGAAATCAGGGAGCTTTATGAAAATGATATTGCTT 1069
QY 943 CCATGA-----GCTCTCACCTGATCCCTGGCCTGAACCTGGCTGCTGTAGGTCTTTTCC 996
Db 1070 CTATGAATCTTCAAGCACAATTTAATCTCTGGATTAATCTGAACGCTTTGGGTCTGTCC 1129
QY 997 CAGCTTCATCCAGCGCAGTCCCGCCCTCCAGCAGGTTACTGGGGCTGCTCCCTATA 1056
Db 1130 CACCCACTTCAGGAGTCCACCTCCACCTCAGGGCCCTTCAGCCATGACTCCTCCCT 1189
QY 1057 GCTCCTTTATGAGGCTCCGAGCAGAGATGTTGAGGTGTTATCCCGCCAGGAG 1116
Db 1190 ACCCGCAGTTTGAACAATCAGAAACGAGACTGTTTCATCTGTTATPCCAGCTCTATCAG 1249
QY 1117 TGGGGCCCATCATCGGCAAGAGGGGAGCAGACATCAACAGCTCTCCCGTTTGCAGCG 1176
Db 1250 TCGGTGCCATCATCCGCAAGCAGGCGCAGCAGCATCAGCAGCTTCTCGCTTTGCTGGAG 1309
QY 1177 CCTCCATCAAGATTGCACCCAGCCGAAACACCTGACTCCAAAGTTCGATGGTTATCATCA 1236
Db 1310 CTTCAATTAAAGATTGCTCCAGCGGAAGCACCAGATGCTAAAGTGAGGATGTTGATATCA 1369
QY 1237 CTGGAACCGCAGAGGCCCAATTCAGGCTCAGGGAAGATCTATGCGAATCTCAAGGAG 1296
Db 1370 CTGGACCAACAGAGGCTCAGTTCAAGGCTCAGGGAAGATTTATGGAAAAATTAAGAAG 1429
QY 1297 AGAAGCTTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTTGCCAGCAT 1356
Db 1430 AAAAGCTTTGTTAGTCTTAAGAGAGAGGTGAAGCTTGAAGCTCATATCAGAGTGCATCCT 1489
QY 1357 CAGCAGCTGGCCGGGTCAATTGGCAAAAGGTGAAAAACGGTGAACGAGTTGCAGAAATTGA 1416
Db 1490 TTGCTGCTGGCAGAGTTATTGGAAAAGGAGGCAAAACGGTGAATGAATTCAGAAATTGT 1549
QY 1417 CGGCACTGAGGTGGTAGTACCAAGAGACCAACCCCTGATGAGAACGACCAAGTTCATCG 1476
Db 1550 CAAAGTCAGAAAGTTGTTGCTCCCTCGTACCAGACACCTGATGAGAAATGACCAAGTGGTTG 1609
QY 1477 TGAATAATCATCGACATTTCTATGTCAGTCAGATGGCTCAACGGAAGATCCGAGACATCC 1536
Db 1610 TCAAAATTAAGTCTGTCATCTGTTGCCAGGTTGCCAGAGAAATTCAGGAATTC 1669
QY 1537 TGGCCCAAGTTAAGCAGCAGCATCAGAAG 1565
Db 1670 TGACTCAGGTAAAGCAGCAGCAACACAG 1698

Search completed: July 24, 2004, 22:28:55
Job time : 843.722 secs


```
QY 523 CCATCCGCAACATCACAAAAAGCCAGCCAGTCCAGATAGAGTGCATAGGAGGAGACG 582
Db 650 CCATTCGGACATCACAAACAGAGCCAGTCTAAATCGATGTCCACGGTAAAGAAAATG 709
QY 583 CAGGTGACGTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGTGTCTCTCCGCTT 642
Db 710 CGGGGGTGTGAGAGTCGATTACTCTCTCTACTCTCTGAGGACCTCTCGGGCTT 769
QY 643 GTAAGATGATCTTGAGATTTATGATTAAGAGGCTAAGGACCAACAAAGGCTCACGAGG 702
Db 770 GTAAGTCTTATCTGGAGATTATGATTAAGAGGCTAAGGACCAACAAAGGCTCACGAGG 829
QY 703 TTCCCTCTGAAGATCTCGCCCATTAATCACTTTGAGGGCGTCTCATTTGGCAGGAGGAC 762
Db 830 TCCCTTGAAGATTTAGCTCATTAATCACTTTGAGGGCGTCTCATTTGGCAGGAGGAA 889
QY 763 GGAACCTGAAGAAGGTAGAGCAAGATACCGAGACAAATAACCATCTCTCTGTTGCAAG 822
Db 890 GAAATCTTAAAAAATTTGAGCAAGACACAGACACTAAATCACGATATCTCCATTGCGAG 949
QY 823 ACCTTACCTTTACAAACCTGAGAGGACCATCATCTGTGAGGGGGCCATCGAATTTGTT 882
Db 950 AATTGACGCTGTAAATCCAGAACGCACTATTACAGTTAAAGGCAATGTTGAGACATGTG 1009
QY 883 GCAGGGCCGAGCAGGAAATTAATGAAGAAAGTTCCGGAGGCGCTATGAGAAATGATGTGCTG 942
Db 1010 CCAAGCTGAGGAGGAGATCATGAGAAATACAGGAGTCTTATGAAATGATATGCTT 1069
QY 943 CATTGA-----GCTCTCACTGATCCCTGCTGAAACCTGGCTGTGTAGTCTTTTCC 996
Db 1070 CTATGATCTTCAAGCACATTTAATCTCTGATTTAAATCTGAACGCTTGGGCTGTCTCC 1129
QY 997 CAGCTTCATCCAGCGAGTCCCGCGCTCCAGCAGCGTACTGGGGCTGTCTCCCTATA 1056
Db 1130 CACCCACTTCAGGATGCCACCTCCACCTCAGGGGCCCTTCAGCCATGACTCTCTCCCT 1189
QY 1057 GCTCTCTTTATGCAAGCTCCCGAGCAGGAGATGTTGCAAGTGTATTATCCCGCCCGAGCAG 1116
Db 1190 ACCCGAGTTTGACATCAGAAACGAGACTGTTCTATCTGTTATCCCGAGCTATACAG 1249
QY 1117 TGGGGCGCATCATCGGCAAGAGGGGAGCAGCATCAAAAGCTCTCTCCGGTTTGCCAGCG 1176
Db 1250 TCGGTGCATCATCGGCAAGAGGGGCCAGCAGCATCAAGCAGCTTCTCTCGCTTGAG 1309
QY 1177 CTTCCATCAAGATTGCAACCGAAGACCTGACTCCAAAGTTCGTATGTTATCATCA 1236
Db 1310 CTTCAATTAGATTGCTCCAGCGAAGCACCAGATGCTAAAGTGAGGATGTTGATTATCA 1369
QY 1237 CTGGACCGCAGAGGCCCAATTCAAGGCTCAGGGAAGAAATCTATGGCAAACTCAAGGAGG 1296
Db 1370 CTGGACCAACAGAGGCTCAGTTCAAGGCTCAGGGAAGAAATTTATGAAATTTAAGAAG 1429
QY 1297 AGAATCTTTGTTCCCAAGGAGGAGTGAAGTGGAGACCCACATAAGTGTGCCAGCAT 1356
Db 1430 AAAAATTTGTTAGTCTTAAAGAGAGGTGAAATTTGAAAGCTCATATCAGAGTGCCATCT 1489
QY 1357 CAGCAGTGGCCGGTTCATTGGCAAGGTGGAAAGCGTGAACGAGTTCAGAAATTTGA 1416
Db 1490 TTGCTGCTGCAGAGTTATTGGAAGAGGAGGCAAAACGGTGAATGAATTCAGAAATTTGT 1549
QY 1417 CGGACGTGAGGTGGTATGTAACCAAGAGACAGACCCCTGATGAGAACGACAGGTATCG 1476
Db 1550 CAAGTGCAAGATTGTTGTCCTCGTGACCAAGACCTGATGAGATGACCAAGTGGTTG 1609
QY 1477 TGAATATCATCGACATTTCTATGCCAGTCAGATGGCTCAACGGAGATCCGAGACATCC 1536
Db 1610 TCAAAATAAATCGTGTCACTTATGCTGTCCAGGTGTCAGAGAAAAATTCAGAAATTC 1669
QY 1537 TGGCCAGAGTTAAGCAGCAGCATCAGNAG 1565
Db 1670 TGACTCAGGTAAGAGCAGCAACCAACAG 1698
```

```
RESULT 15
US-09-897-778-447
; Sequence 447, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrik
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121-455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 447
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-897-778-447
```

```
Query Match 34.6%; Score 674.2; DB 9; Length 1743;
Best Local Similarity 66.7%; Pred. No. 6.6e-200;
Matches 980; Conservative 0; Mismatches 483; Indels 6; Gaps 1;
```

```
QY 103 GGGTTAGGAGCCGGAAAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTAC 162
Db 230 GGCAAGGATTCGAAACTTCAGATACGAAATATCCGCCTCATTTACAGTGGAGTGC 289
QY 163 TGCACAGCTGCTGGCTCAGTATGTTACAGTACAGAACTCTGAGCAAGTGAACACCGAGA 222
Db 290 TGGATAGTTTACTAGTCCAGTATGGAGTGTGAGAGCTGTGAGCAAGTGAACACTGACT 349
QY 223 GTGAGACGGCAGTGTGAATGTCACTATTCACCCGAGCAGACCCAGCAAGCCATCA 282
Db 350 CGSAAACTGAGTTGTAATGTAACCTATTCAGTAAAGCAAGCTAGCAAGCACTAG 409
QY 283 TGAAGCTGATGCCCACAGTGTGGAGAACCATGCCCCCTGAAGTCTCTCATCCCCGATG 342
Db 410 ACAAACTGAATGGATTTTCAGTTAGAGAAATTCACCTTGAAGTAGCTATATCCCTGATG 469
QY 343 AGCAGATAGCACAGGACCTGAGAAATGGGGCCGAGGGGCTTTGGCTCTCGGGGTACG 402
Db 470 AAACGGCCGACAGAAACCCCTTGACAGCCCCGAGGTGCCCCGGGGCTTTGGGCGAG 529
QY 403 CCGCCAGGGCTCACCTGTGGCAGCGGGGGCCCCAGCCAGCAGCAGCAAGTGGACATCC 462
Db 530 GGGGCTCTCAAGGCAAGGCTCTCCAGGATCCGTATCCAAAGCAGAACCATGTGATTTC 589
QY 463 CCCTTCGGCTCTCGTGGTCCCAACCCAGTATGTTGGGTGCCATTTATGTCAGAGGGGGCCA 522
Db 590 CTCTGGCCCTGCTGGTTCCTCCACCAATTTGTTGGAGCCATCATAGGAAAGAGGTGCCA 649
QY 523 CCATCCGCAACATCACAAACAGACCCAGTCCAAAGATAGACGTGCTATAGAGGAGGAGCG 582
Db 650 CCATTCGGAAACATCACAAACAGACCCAGTCTAAATCGATGTCCACCCTTAAGAAAAATG 709
QY 583 CAGGTSAGCTGAAAAAGCCATCAGTGTGCACTCCACCCCTCAGGGCTCTCTCTCCGCTT 642
Db 710 CGGGGCTCTCTGAGAAGTCTGATTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 769
QY 643 GTAAGATGATCTTGAGATTTATGATTAAGAGGCTAAGGACCAACAAAGGCTGAGAGG 702
Db 770 GTAAGTCTTATCTGGAGATTATGATTAAGAGGCTAAGGACCAACAAATTCAGAAAGAGA 829
QY 703 TTCCCTCTGAAGATCTTGGCCCATTAATCACTTTTGGGGCTCTCATTTGCCAAGGAGGAC 762
```



```
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-117-982-478

Query Match      34.6%; Score 674.2; DB 15; Length 1740;
Best Local Similarity 66.7%; Pred. No. 6.6e-200;
Matches 980; Conservative 0; Mismatches 483; Indels 6; Gaps 1;

QY 103 GGGTTAGGACCGGAAATTCAAATCCGAAATATTCACCCAGCTCCCGATGGGAAGTAC 162
Db 230 GGCAAGGATTCGGAACTTCAGATACGAAATATCCCGCTCATTTACAGTGGAGGTGC 289

QY 163 TGGACAGCTGTGGCTCAGTATGGTACAGTAGAGAACTGTGAGCAAGTGAACACGAGA 222
Db 290 TGGATAGTTTACTAGTCCAGTATGGAGTGTGGAGAGCTGTGAGCAAGTGAACACTGACT 349

QY 223 GTGAGACGGCAGTGGTGAATGTCACTATTCCAAACGGGAGCAGACCCAGGCAAGCCATCA 282
Db 350 CGGAACTGCAGTTGTAATGTAACTTCCAGTAAGGACCAAGCTAGACAGCACTAG 409

QY 283 TGAAGTGAATGGCCACCACTGTGAGAACCATGCCCTGAAGGTCTCTTACATCCCCGATG 342
Db 410 ACAAACTGAATGGATTTCAGTTAGAGAAATTCACCTTTGAAAGTAGCCTATATCCCTGATG 469

QY 343 AGCAGATAGCACAGGACCTGAGNAATGGGCGCGAGGGGCTTTGGCTCTCGGGGTGAGC 402
Db 470 AAGCGCGCGCCAGCAAAACCCCTTGCAGCAGCCCCGAGGTCGCGGGGGCTTTGGGCAGA 529

QY 403 CCGGCCAGGGCTCACCTGTGGCAGCGGGGGCCCCAGCAAGCAGCAGCAAGTGCACATCC 462
Db 530 GGGGCTCTCAAGGCAGGGGTCTCCAGGATCCGATATCCAAGCAGAAACCATGTGATTTGC 589

QY 463 CCCTTGGCTCTGGTGCCCAACCAAGTATGGGTGCCATTAATTGGCAAGGAGGGGCCA 522
Db 590 CTCTGGCTCTGGTGTCCACCCCAATTTGTTGGGCCATCATAGGAAAGAAAGGTGCCA 649

QY 523 CCATCCGACATCATCAAAACACAGCCAGTCCAGTAGACGTGCTATAGGAAGAGAACG 582
Db 650 CCATTCGGAACATCACCAACACAGCCAGTCTAAATTCGATGTCACCGTAAAGAAATG 709

QY 583 CAGGTGAGCTGAAAGCCATCAGTGTGACCTCCACCCCTGAGGGTGTCTCTCCGCTT 642
Db 710 CGGGGCTGTGAGAGTCCGATTACTATCTCTCTACTCTCTGAGGACCTCTGCGGCTT 769

QY 643 GTAAGATGATCTTGGAGATTATGATTAAGAGGCTAAGGACACCAAAACGGCTGACGAG 702
Db 770 GTAAGTCTATTCTGGAGATTATGCATAAGGAAGCTCAAGATATAAAATTCACAGAAAGA 829

QY 703 TTCCCTGGAAGACCTGGGCCATAATAAATTTGAGGGCTCTCATTTGGCAAGGAGGAC 762
Db 830 TCCCTTGAAGATTTTAGCTCAATAAATCTTTGAGACCTCTATTGGTAAAGAAAGGAA 889

QY 763 GGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAATTCACCATCTCTCTCTGTCGAG 822
Db 890 GAAATCTTAAATAATTGACNAGACACACACACTAAATACAGATATCTCCATTGCAAG 949

QY 823 ACCTTACCTTTACACCTTGAGAGGACCAATCACTGTGAAGGGGGCCATCGAATTTGTT 882
Db 950 AATTGACGTGTATAATCCAGAACCGCACTATTACAGTTAAAGCAATGTTGAGACATGTG 1009

QY 883 GCAGGGCCGAGCGGAATAATCAAGAAAGTTCCGGAGGCTCATGAGATGATGGCTG 942
Db 1010 CCAAGCTGAGGAGGAGATCATGAAGAAATCAGGGAGTCTTTATGAAATGATATTGCTT 1069

QY 943 CCATGA-----GCTCTCACTGATCCCTGGCTGAACCTGGCTGTGTAGTGTCTTTCC 996
Db 1070 CTATGAATCTTCAAGCACATTTAATCTCTGATTAATCTGAAGGCTTGGGCTGTGTTCC 1129

QY 997 CAGCTTCATCCAGGCAAGTCCCGCGCTCCAGCAGCGGTTACTGGGGTGTCTCCCTATA 1056
Db 1130 CACCCACTTCAGGGATGCCACCTCCACCTCAGGGGCCCCCTTCAGGCCATGACTCTCCCT 1189
```

```
QY 1057 GCTCCTTTATGCAAGGTCCCGACGAGAGATGCTGCAAGTGTTTATCCCGCCCGACGAG 1116
Db 1190 ACCCGAGATTGAGCAATCAGAAACGAGAGACTGTCATCTGTTTATCCAGCTCTATCAG 1249

QY 1117 TGGGCGCCATCATCGGCAAGAGGGGCGACACATCAAAAGCTCTCCCGTTTGGCAGCG 1176
Db 1250 TCGGTGCAATCATCGGCAAGCAGGGCCGACACATCAAGCAGCTTCTCGCTTTGCTGGAG 1309

QY 1177 CTTCCATCAAGATTGCAACACCCGAAACCTGACTCTCAAAAGTTCTGATGTTATCATCA 1236
Db 1310 CTTCAATTAAGATTGCTCCAGCGAAGCACAGATGCTAAAGTCAGGATGTTGATATCA 1369

QY 1237 CTGACCGCCAGAGGCCAATTCAGGCTCAGGGAAGATCTATATGGCAAACTCAAGAGG 1296
Db 1370 CTGACACCCAGAGGCTCAGTTCAAGGCTCAGGGAAGATTTATGGAATAATTAAGAAG 1429

QY 1297 AGAACTTCTTTGCTCCCAAGGGAAGTGAAGCTGGAGACCCACATACGTTGTCAGCAT 1356
Db 1430 AAACTTTGTTAGTCTTAAAGAGAGGTGAATCTGAAGCTCATATCAGAGTGCCATCCT 1489

QY 1357 CAGCAGCTGCGCGGTCATTGGCAAAAGGTGGAAAAACGGTGAACGAGTTTCAGAAATTTGA 1416
Db 1490 TTGCTGCTGCGCAGAGTTATTGGAAAAAGGAGGCAAAACGGTGAATGAATTCAGAAATTTGT 1549

QY 1417 CGCAGCTGAGGTGTAGTATCAAGAGACAGACCCCTGTATGAGAACGACGAGGTCTATCG 1476
Db 1550 CAAAGTCGAGAAAGTTGTTGTCCTCGTGCACAGACACCTGATGAGAATGACCAAGTGGTTG 1609

QY 1477 TGAATAATCATCGGACATTTCTATGCCAGTCAGATGCTCAACGGAAGATCCGAGACATCC 1536
Db 1610 TCAAAATTAAGTGTCTCTCTCTGCTTGCAGGTTGCCAGAGTTCAGGAAAAATTCAGGAAATTC 1669
```

RESULT 13

```
US-10-313-986-347
; Sequence 347, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-313-986-347
```

```
Query Match      34.6%; Score 674.2; DB 15; Length 1740;
Best Local Similarity 66.7%; Pred. No. 6.6e-200;
Matches 980; Conservative 0; Mismatches 483; Indels 6; Gaps 1;
```

```
QY 103 GGGTTAGGACCGGAAATTCAAATCCGAAATATTCACCCAGCTCCCGATGGGAAGTAC 162
Db 230 GGCAAGGATTCGGAACTTCAGATACGAAATATCCCGCTCATTTACAGTGGAGGTGC 289

QY 163 TGGACAGCTGTGGCTCAGTATGGTACAGTAGAGAACTGTGAGCAAGTGAACACGAGA 222
Db 290 TGGATAGTTTACTAGTCCAGTATGGAGTGTGGAGAGCTGTGAGCAAGTGAACACTGACT 349

QY 223 GTGAGACGGCAGTGGTGAATGTCACTATTCCAAACGGGAGCAGACCCAGGCAAGCCATCA 282
```

; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-117-982-347

Query Match 34.6%; Score 674.2; DB 15; Length 1740;

Best Local Similarity 66.7%; Pred. No. 6.6e-200;

Matches 980; Conservative 0; Mismatches 483; Indels 6; Gaps 1;

Qy	103	GGGTTAGGAGCGGAAATTCGAAATATTCACACCCAGCTCCGATGGGAAGTAC	162
Db	230	GGCAAGGATTTCGAAATCTTCAGATACGAAATATCCCGCTCATTTACAGTGGAGTGC	289
Qy	163	TGGACAGCTCTGCTCAGTATGGTACAGTAGAGAACTGTGAGCAAGTGAACCCGAGA	222
Db	290	TGGATAGTTTACTAGTCAGTATGGAGTGGTGGAGAGCTGTGAGCAAGTGAACACTGACT	349
Qy	223	GTGACAGCGAGTGGTGAATGTACCTATTCACACCGGGAGGAGCAGCAGGCAAGCCATCA	282
Db	350	CGGAACTGCGAGTTGTAATTAACCTATTCCAGTAAAGGACCAAGCTAGACAAAGCACTAG	409
Qy	283	TGAAGCTGAATGGCCACAGTTGGAGAACCATGCCCTGAAGTCTCCTACATCCCGCATG	342
Db	410	ACAACTGAATGGATTTCAGTTAGAGAAATTCACCTTGAAAGTAGCCTATATCCCTGATG	469
Qy	343	AGCAGATAGCAGGAGCCTGAGAATGGGCGCGAGGGGGCTTTGGCTCTCGGGGTACGC	402
Db	470	AAACGGCGCCAGCAAAACCCCTTGCAGCAGCCCGAGGTGCGCGGGGCTTGGGCAGA	529
Qy	403	CCGCGCAGGGTTCACCTGTGGCAGCGGGGGCCCGCAGCAAGCAGCAGCAGTGGACATCC	462
Db	530	GGGGCTCTCAGGCGAGGGGTCTCCAGGATCCGATATCCAAAGCAGAAACCATGTGATTTC	589
Qy	463	CCCTTCGGCTCTGTCGCCACCCAGTATGGGTGCCATTATTCGCAAGGAGGGGCCA	522
Db	590	CTCTGCGCTCTGCTTCCCACCCAAATTTGTTGGAGCCATCATAGAAAGAGAGTGCCA	649
Qy	523	CCATCCGCAACATCAAAAAACAGCCAGTCCAGATAGACGCTGCATAGGAAGGAGAAG	582
Db	650	CCATTCGGAAATCATCACCAACAGACCCAGTCTAAATCGATGCCACCGTAAAGAAATG	709
Qy	583	CAGTGCAGCTGAAAGGCATCAGTGTGCATCCACCCCTGAGGCGTCTCTCGGCTT	642
Db	710	CGGGGCTGCTGAGAAGTCGATTACTATCTCTACTCTGAAAGGCACCTCTCGGGCTT	769
Qy	643	GTAAGTATCTTTCGAGATTATGCAATAAGAGGCTTAAGGACACCAAAACGGCTGACGAG	702
Db	770	GTAAGTCTATTCGAGATTATGCAATAGGAAGCTCAAGATATAAAATTCAGAGAGA	829
Qy	703	TTCCCTGAAGATCTCGGCCCAATAACTTTGTAGGGCGTCTCATTTGGCAAGGAGGAC	762
Db	830	TCCCTTGAAGATTTTAGTCTATAATACTTTTGGAGCTCTTATTTGTAAGAAGGAA	889
Qy	763	GGAACCTGAGAGGTAGCAGAGATACCGACACCAAAATCAACCTCTCTCTGTCAG	822
Db	890	GAAATCTTAAATAAATTTAGCAGAGACACAGACACTAAATACGATATCTCCATTCAGG	949
Qy	823	ACCTTACCTTTTACAACTCTGAGAGACCATCACTGTGAAAGGGGGCCATCGAGAAATGTT	882
Db	950	AATTGACGCTGTATATCCAGAACGCACTATTACAGTTAAAGCAATGTTGACACATGTG	1009

Qy	883	GCAGGCGGAGCAGGAATATGAAGAAAGTTTCGGAGGCCTATGAGAATGATGGCTG	942
Db	1010	CCAAAGCTGAGGAGGAGATCATGAAGAAATCAGGAGTCTTATGAATAATATGCTT	1069
Qy	943	CCATGA-----GCTCTCACCTGATCCCTGGCCTGAACCTGGCTGCTAGTCTTTTCC	996
Db	1070	CTATGAATCTCAAGCACAATTAATTCCTGGATTAAATCTGAACGCTTGGGTCTGTTCC	1129
Qy	997	CAGCTTCATCCAGCGCAGTCCCGCGCTCCAGCAGGTTACTTGGGCTGCTCCCTATA	1056
Db	1130	CACCACTTCAGGAGTCCACTCCACCTCAGGCGCCCTTCAGCCAATGACTCTCCTCT	1189
Qy	1057	GCTCCTTTATGAGGCTCCGAGCAGGAGATGTGAGTGTATTATCCCGCCAGGAG	1116
Db	1190	ACCGCAGTTTGAGCAATCAGAAACGGAGACTGTTCTATCTGTTATCCAGCTCTATCAG	1249
Qy	1117	TGGCGCCATCATCGCAAGAGGGGAGCAGCACAACACAGCTCTCCCGTTTCCAGCG	1176
Db	1250	TCGTTGCCATCATCGCAAGCAGGCGCAGCACAATCAAGCAGCTTCTCGCTTGGAG	1309
Qy	1177	CCTCCATCAAGATTGCACCCCGAAACACCTGACTCCAAAGTTCGATGTTATCATCA	1236
Db	1310	CTTCAATTAAGATTGCTCCAGCGGAAGCAGAGATGCTTAAAGTGAAGTGGATTATCA	1369
Qy	1237	CTGACCCCGCAGAGCCCAATTCAGGCTCAGGGAAGATCTATGSCAAACTCAAGGAGG	1296
Db	1370	CTGACCAACAGAGGCTCAGTTCAAGGCTCAGGGAAGAAATTTATGAAAAAATTAAGAG	1429
Qy	1297	AGAACTCTTTCGTCCTCCAAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGAT	1356
Db	1430	AAAACTTTGTTAGTCTTAAAGAGAGGTGAACCTTGAAGCTCATATCAGAGTGCCTCT	1489
Qy	1357	CAGCAGCTGGCGGGTCAATTTGGCAAAAGGTGAAAAACCGTGAACAGTTCAGAAATTTGA	1416
Db	1490	TTGCTGCTGGCAGAGTTATTGAAAAAGGAGGCAAAACCGTGAATGAACCTTCAGAAATTTGT	1549
Qy	1417	CGCAGCTGAGGTGTAGTACCAAGACCAAGACCCCTGATGAGAACCAAGAGTGCATCG	1476
Db	1550	CAAGTGCAGAGTTGTTGTCCTCGTGACCAAGACCTGATGAGTAATGACCAAGTGTG	1609
Qy	1477	TGAAATCATCGACATTTCTATGCGAGTTCAGATGGCTCAAGCGAAGATCCGAGACATCC	1536
Db	1610	TCAAAATACCTGGTCACTTCTATGCTTCCAGGTTGCCAGAGAAAAATTCAGGAAATTC	1669
Qy	1537	TGGCCAGGTTAAGCAGCAGCATCAAG	1565
Db	1670	TGACTCAGTAAAGCAGCACCACCAACAG	1698

RESULT 12

US-10-117-982-478

; Sequence 478, Application US/10117982

; Publication No. US20030138438A1

; GENERAL INFORMATION:

; APPLICANT: Foy, Teresa M.

; APPLICANT: Fanger, Gary R.

; APPLICANT: Vedwick, Thomas S.

; APPLICANT: Carter, Darick

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Henderson, Robert A.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Mericle, Barbara

; APPLICANT: Spies, Gregory A.

; APPLICANT: Fan, Liqun

; APPLICANT: Wang, Tongtong

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.455C18

; CURRENT APPLICATION NUMBER: US/10/117,982

; CURRENT FILING DATE: 2002-04-05

; NUMBER OF SEQ ID NOS: 484

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 478

APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Li, Samuel X.
APPLICANT: Kalos, Michael D.
APPLICANT: Henderson, Robert A.
APPLICANT: McNeill, Patricia D.
APPLICANT: Fanger, Neil
APPLICANT: Retter, Marc W.
APPLICANT: Durham, Margarita
APPLICANT: Fanger, Gary R.
APPLICANT: Vedwick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Watanabe, Yoshihiro
APPLICANT: Peckman, David W.
APPLICANT: Cai, Feng
APPLICANT: Foy, Teresa M.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.455C17
CURRENT APPLICATION NUMBER: US/10/007,700
NUMBER OF SEQ ID NOS: 469
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 347
LENGTH: 1740
TYPE: DNA
ORGANISM: Homo sapiens
US-10-007-700-347

Query Match 34.6%; Score 674.2; DB 13; Length 1740;
Best Local Similarity 66.7%; Pred. No. 6.6e-200;
Matches 980; Conservative 0; Mismatches 483; Indels 6; Gaps 1;

QY	103	GGGTTAGGAGCGGAAATTC	AAATTCGAAATTTCCACCC	AGCTCCGATGGGAAGTAC	162
Db	230	GGCAAGGATTCGAAACCTT	CAGATACGAAATATCCCGC	CTCATTTACAGTGGGAGTGC	289
QY	163	TGCACAGCTGCTGGCTCAG	TATGTTACAGTACGAGACT	GTGAGCAAGTGAACACCGAGA	222
Db	290	TGATAGTTTACTAGTCCAG	TATGAGTGTGGAGAGCTGT	GAGCAAGTGAACACTGACT	349
QY	223	GTGAGACGGCAGTGTGAAT	GTGCACTTATCCACCGGAG	CAGACAGCAGCAAGCCATCA	282
Db	350	CGAAACTGCAGTTGAATGT	AACTTCCAGTAAAGACCA	AGCTAGACAGCACTAG	409
QY	283	TGAGCTGATGCGCACCGAT	TGGAGAACATCCCTGAAAG	GTCTCTTACATCCCGATG	342
Db	410	ACAACTGAATGGAATTCAG	TTAGATTAGAAATTTCACT	TGAAAGTAGCTATATCCCT	469
QY	343	AGCAGATAGCACAGGACCT	GAGATGGGCGCGAGGGGCT	TTGGCTCTCGGGGTGAGC	402
Db	470	AAACGGCCCGCAGCAAAAC	CCCTTGCAGCAGCCCGAGT	CGCCGGGGCTTGGGCAGA	529
QY	403	CCGCGACGGGCTCACTGTG	GCAGCGGGGCCCCAGCCAG	CAGCAGCAAGTGGACATCC	462
Db	530	GGGGCTCCTCAGGCGAGGG	GTCTCCAGGATCCGTATCC	AAAGAGAACCATGTGATTGC	589
QY	463	CCCTTCGGCTCTGGTGGCC	ACACCGATGTGGGTGCCATT	TGCAAGAGGGGGGCCA	522
Db	590	CTCTGGCTGTGGTGTCCAC	CCCAATTTGTGGAGCCATC	ATAGGAAAGAGGTGCCA	649
QY	523	CCATCCGCAACATCACAAAC	AGACCCAGTCCAAAGATAG	ACGTGCATAGAGGAGAAGC	582
Db	650	CCATTCCGAACATCACCAAC	AGACCCAGTCTAAATCGAT	TGCCACCGTAAGAAATG	709
QY	583	CAGTGCAGTGAAGCAATCAG	TGTGCTACCTCCACCCCTG	AGGGGCTGTCTCCGCTT	642
Db	710	CGGGGCTGTGAGAGTCGAT	TACTCTCTTACTCTCTGAG	GCACCTCTGCGGCTT	769
QY	643	GTAAGATGATCTGGAGATTG	CATTAAGAGGCTAAGACAC	CAACCAAAACGGCTGACAGG	702
Db	770	GTAAGTCTATTCTGGAGATT	ATGATGAAGAGCTCAAGAT	ATAAAATTCACAGAGAGA	829

RESULT 11

US-10-117-982-347
; Sequence 347, Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.

QY	703	TTCCCTTGAAAGATCCTGG	CCCCATAATAAATTGTTAG	GGCGCTCTCATTTGGCAAG	GAAGAC	762
Db	830	TCCTCTTGAAAGATTTAG	CTCATAATAAATTGTTG	GACGCTTATTTGGTAAAG	AGGAA	889
QY	763	GGAACTTGAAAGAGTAGA	CAAGATACCGAGACAAAT	ATCACCATCTCTCTGTC	CAAG	822
Db	890	GAAATCTTAAATAAATTT	GAGCAAGACACAGACTAA	ATATTCAGATATCTCCAT	TTCAGG	949
QY	823	ACCTTACCTTTACAACCC	CTGAGAGGACCATCACTG	TGAAGGGGGCCATCAGA	AAATGTT	882
Db	950	AATTGAGCTGTATATCC	AGAACGCACTATTACAG	TAAAGSCAATGTTGACA	CATGTG	1009
QY	883	CGAGGCCGAGCAGGAAAT	TAATGAAGAGTTCCGGAG	GCCTATGAGATGATGTG	GGCTG	942
Db	1010	CCAAAGCTGAGGAGGATC	ATGAAGAAATCAGGGAGT	CTTATGAAAATGATATTG	CTT	1069
QY	943	CCATGA-----GCTCT	CACTGATCCCTGGCCCTG	CACTGCTGCTGAGGCTTT	TC	936
Db	1070	CTATGATCTTCAAGCACA	TTAATTCCTGGATTAATC	TGAACGCTTGGGCTGTG	TTCC	1129
QY	997	CAGCTTCATCCAGCGCAG	TCCCGCCCTCCAGCAGCG	TTACTGGGGCTGCTCCCT	TATA	1056
Db	1130	CACCCACTTCAGGGATGC	CACTCCACCTCAGGGCCCC	CTTCAGCCATGACTCTCT	CCCT	1189
QY	1057	GCTCCTTTATGAGGCTCC	CGAGCAGGATGCTGAGGT	TTTATTCCTCCGCCAGG	CGAG	1116
Db	1190	ACCCGAGTTTGAGCAAT	CAGAAACGGAGACTGTT	CTCATCTGTTATCCCGA	CTTATCAG	1249
QY	1117	TGGCGCCCATCATCGGCA	AGAGGGGCGACACATCAA	ACAGCTCTCCCGTTTGC	CAGCG	1176
Db	1250	TCGCTGCCATCATCGCA	AGCAGGGCCAGCACAAT	CAAGCAGCTTCTCGCT	TTCGGAG	1309
QY	1177	CTTCATCAAGATTCACCA	CCCGAAACCTGACTCCAA	AGTTCTGATTTATCATCA		1236
Db	1310	CTTCAATTAGATTTGTC	CCAGCGAAGCACAGATCT	CTAAAGTGAGGATGGTG	ATTATCA	1369
QY	1237	CTGACCGCAGAGGCCCAT	TTCAGGCTCAGGAGATCT	ATGCGAACTCAAGGAGG		1296
Db	1370	CTGGACCAACAGAGGCTC	AGGCTCAGGAAAGAAAT	TATGGAATAATTTAAAG	AAAG	1429
QY	1297	AGAACTTCTTTGCTCCCA	AGGAGGAGTGAAGCTG	GAGACCCACATACGTG	CCAGCAT	1356
Db	1430	AAACTTTCTTAGTCTTAA	AGAGAGTGAACTGAACT	CTATATCAAGTGCCTCT		1489
QY	1357	CAGAGCTGGCGGGTCATT	GGCAAGGTGGAATAACCG	TGAACGAGTTTCAGAA	TTTGA	1416
Db	1490	TTGCTCTGGCAGAGTTAT	TGGAAAGAGGAGCAAA	ACCGTGAATGAATTCAG	AAATTTGT	1549
QY	1417	CGGCACTCAGTGGTGTAG	TACCAAGAGACGACCC	CTGATGAGACGACCATCG		1476
Db	1550	CAAGTCAGAAAGTTGTTG	CTCCCTCGTGACACACCT	GATGAGAAATGACCAAG	TGGTTG	1609
QY	1477	TGAAATCATCGGACATTT	CTATGCGCAGTCAAGTGC	TCAACGGAAGATCCGAG	ACATCC	1536
Db	1610	TCAAAATAACTGTCAC	TTCTATGCTTGCCAGGT	TGCCAGAGAAAATTCAG	AAATTC	1669
QY	1537	TGCCCCAGGTTAAGCAG	CGAGCATCGAAG			1565
Db	1670	TGACTCAGGTAAGCAGC	ACCAACACAG			1698

Db 1310 CTTCAATTAAAGATTGCTCCAGCGGAGCACAGATGCTAAAGTGAGGATGGTATTATCA 1369
QY 1237 CTGAGACCGCAGAGGCCAATTCAAGGCTCAGGGAAGAAATCTATGGCAAACTCAAGGAGG 1296
Db 1370 CTGGACACACAGAGGCTCAGTTCAGGCTCAGGGAAGAAATTTATGGAAAAATTAAGAAG 1429
QY 1297 AGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCATACATAGTGTGCAGCAT 1356
Db 1430 AAAACTTTGTAGTCTTAAGAAGAGGTGAACCTTGAAGCTCATATCAGATGCGCATCCT 1489
QY 1357 CAGCAGCTGCGCGGTCATTGGCAAAAGTGGAAAAACGGTGAACGAGTTCAGAAATTTGA 1416
Db 1490 TTGCTGCTGCGAGAGTTATTGGAAAAAGGAGCAAAACGGTGAAGTTCAGAAATTTGT 1549
QY 1417 CGCAGCTGAGGTGGTAGTACCAAGAGACACAGACCCCTGATGAGAAACGACAGGTCAATCG 1476
Db 1550 CAAAGTCAGAAAGTTGTTGCTCCCTGTCAGCAGACACCTGATGAGAAATGACCAAGTGGTTG 1609
QY 1477 TGAATAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCGAGACATCC 1536
Db 1610 TCAATAATCACTGTCATCTATGCTTGCAGGTTGCCAGAGTGGCCAGAAAAATTCAGGAAATTC 1669
QY 1537 TGGCCCAAGGTTAAGCAGCAGCATCAGAAAG 1565
Db 1670 TGACTCAGGTAAAGCAGCAGCAACCAACAG 1698

RESULT 8

US-09-850-716A-347
; Sequence 347, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-850-716A-347

Query Match 34.6%; Score 674.2; DB 9; Length 1740;
Best Local Similarity 66.7%; Pred. No. 6.6e-200;
Matches 980; Conservative 0; Mismatches 483; Indels 6; Gaps 1;
QY 103 GGGTTAGGAGCGGAAATTCAAATCCGAAATATCCACCCAGCTCCGATGGGAAGTAC 162
Db 230 GGCAAGGATTCGAAACCTTCAGATACGAAATATCCCGCTCATTTACAGTGGGAGTGC 289
QY 163 TGGACAGCTGCTGGCTCAGTATGTACAGTAGAGAACTGTGAGCAAGTGAACACCGAGA 222
Db 290 TGGATAGTTTACTAGTCCAGTATGGAGTGTGGAGAGCTGTGAGCAAGTGAACACTGACT 349
QY 223 GTGAGACGGCAGTGTGAATGTCACTATTCCAAACCGGAGCAGACCGAGGCCATCA 282
Db 350 CGGAAACTGCAGTTGTAATGTAACCTATTCAGTAGAGACCCAGCTAGACAGCACTAG 409
QY 283 TGAAGCTGAATGCCACAGATTGGAGAACATATCCCTGAAGTCTCTATACATCCCGATG 342
Db 410 ACAAACTGAATGGATTTCAATTAGAGAAATTCACCTTGAAGTAGCCTATATCCCTGATG 469
QY 343 AGCAGATAGCAGAGGACCTGAGAAATGGGCGCAGGGGCTTTGGCTCTCGGGGTACGC 402
Db 470 AAACGCCGCCAGCAAAACCCCTTTCAGAGCAGCCCCGAGGTGCGCGGGGGCTTTGGGCAGA 529

QY 403 CCCGCCAGGGGCTCACTGTGGCAGCGGGGCCCCAGCCAGCAGCAGCAAGCTGGACATCC 462
Db 530 GGGGCTCCTCAAGGCAAGGGGTCTCCAGGATCCGATATCCAAGCAGAAACCATGTGATTGC 589
QY 463 CCCTTGGGCTCCTGGTCCACCCAGTATGTGGTGGCCATATTGGCAAGAGGGGGCCA 522
Db 590 CTCTGCGCCCTGCTGGTTCCTCCCAACCCCAATTTGTGGAGCCATCATAGGAAAAAGAGTCCCA 649
QY 523 CCATCCGCAACATCACAAAAACAGACCCAGTCCAAAGATAGACGTGCAATAGGAAGAGAACG 582
Db 650 CCATTGGGAACATCACCAACAGACCCAGTCTAAATCGATGTCCACCGTAAAGAAATG 709
QY 583 CAGGTGACGTGAAAAGCCATCAGTGTGCCTCCACCCCTGAGGGGTGTCTCTCCGCTT 642
Db 710 CGGGGGCTGCTGAGAAGTCGATTACTATCTCTCTACTCTGGAAGCACCCTCTCGGGCTT 769
QY 643 GTAAGATGATCTTGGAGATTATGATAAAGAGCTAAGGACCAAAACCGCTGACAGG 702
Db 770 GTAAGTCTATCTGGAGATTATGATAAGAGCTCAAGATATAAAATTCACAGAGAGA 829
QY 703 TTCCCTGAAGATCCTGGCCCATATAAATCTTTGTAGGGCTCTCATTTGGCAAGGAAGGAC 762
Db 830 TCCCTTGAAGATTTTAGCTCATATAACTTTGTGACGTCTTATTGGTAAAGAGGAA 889
QY 763 GGAACCTGAAGGTTAGACAAAGATACCGAGACAAAATTCACCATCTCTCTCTGCAAG 882
Db 890 GAAATCTTAAAAAAATTTGAGCAAGACACACACTAAATTCAGATATCTCCATTGCGAG 949
QY 823 ACCTTACCCCTTTACAACCCCTGAGAGACCATCACTGTGAAGGGGCCATCGAGAATTTGT 882
Db 950 AATTGAGCTGTATATTCAGAAAGCAGCATTTACAGTTAAAGCAATTTGTGACATGTG 1009
QY 883 GCAGGGCCGAGCAGGAAATAATGAAGAAAGTTCCGGAGGCCTATGAGAATGATGTGGCTG 942
Db 1010 CCAAAGCTGAGGAGGAGATCATGAAGAAATCAGGGAGTCTTATGAAATGATATTGCTT 1069
QY 943 CCAATGA-----GCTCTCACTGATCCCTGGCTGAACTGGCTGTGTAGTCTTTTCC 996
Db 1070 CTATGAATCTTCAAGCACATTTAAATCTCTGGATTTAAATCTGAACGCTTTGGGTCTGTCC 1129
QY 997 CAGCTTTCATCCAGCGCAGTCCCGCCCTCCAGCAGCGGTTACTGGGGCTGCTCCCTATA 1056
Db 1130 CACCCACTTCAGGATGCCACTCCACCTCAGGGCCCTTCAGCCATGACTCTCCCT 1189
QY 1057 GCTCTTTATGAGGCTCCCGAGCAGGATGATGTCAGGTGTTTATCCCGCCAGGAG 1116
Db 1190 ACCCGCAGTTTGAGCAATCAGAAAACGGAGACTGTTCTATCTGTTTATCCAGCTCTATCAG 1249
QY 1117 TGGGGCCATCATCGCAAGAGGGGCGACACATCAACAGCTCTCCCGTTTGCAGCG 1176
Db 1250 TCGTGCCATCATCGCAAGCGGGCCAGCAGCATCAAGCAGCTTCTCTCGCTTTGCTGGAG 1309
QY 1177 CTTCCATCAAGATTGCAACCCCGAAACACCTGACTCCAAAGTTCGTATGGTTATCATCA 1236
Db 1310 CTTCAATTAAAGATTGCTCCAGCGGAAGCACCAGATGCTTAAAGTGAGGATGGTATTATCA 1369
QY 1237 CTGAGACCGCCAGAGGCCCAATTCAGAGTTCAGGGAAGAAATCTATGGCAAACTCAAGGAG 1296
Db 1370 CTGGACCAACAGAGGCTCAGTTCAAGGCTCAGGGAAGAAATTTATGAAAAATTAAGAAG 1429
QY 1297 AGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCAT 1356
Db 1430 AAAACTTTGTAGTCTTAAAGAGAGGTTGAACCTTGAAGCTCATATCAGAGTGCCATCCT 1489
QY 1357 CAGCAGCTGGCCGGTCTATTGGCAAAAGTGGAAAAACGGTGAACGAGTTGCAAGAAATTTGA 1416
Db 1490 TTGCTGCTGGCAGAGTTATTGGAAAAAGGAGCAAAACGGTGAATGACTTCAGAAATTTGT 1549
QY 1417 CGGCAAGTGAAGTGTGATGTAACAGAGACAGACCCCTGATGAGAACGACCAAGGTCACTG 1476
Db 1550 CAAAGTGCAGAGTTGTTGTCTCTGTCGACACACCTGATGAGAAATGACCAAGTGGTTG 1609
QY 1477 TGAATAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCGGAGACATCC 1536

541 AAAACGGTGAACGAGTTGCAAGATTGACCGGAGCTGAGTGGTAGTACCAAGAGACCAG 600
1449 ACCCTGATGAGAACGACGAGGTCACTCGTGAATAATCATCGACATTTCTATGCGAGTCAG 1508
601 ACCCTGATGAGAACGACGAGGTCACTCGTGAATAATCATCGACATTTCTATGCGAGTCAG 660
1509 ATGGCTCAACGGGAAGATCCGAGACATCCTGGCCAGGTTAAGCAGAGCATCAGAAGGGA 1568
661 ATGGCTCAACGGGAAGATCCGAGACATCCTGGCCAGGTTAAGCAGAGCATCAGAAGGGA 720
1569 CAGAGTAACAGGCGCCAGGACGAGAGTGAACGAGGCTCCCTGTCCTTTGAGTCC 1628
721 CAGAGTAACAGGCGCCAGGACGAGAGTGAACGAGGCTCCCTGTCCTTTGAGTCC 780
1629 AGGACAACAACGGGACGAGAGTGTCTCTCCCGGAGGCTTGAGATTTGAGTGG 1688
781 AGGACAACAACGGGACGAGAGTGTCTCTCCCGGAGGCTTGAGATTTGAGTGG 840
1689 GAATCCGGGACACNTGGGCGGGCTGTAGATCAGGTTTGGCCCACTTGAGTGAAGATG 1748
841 GAATCCGGGACACNTGGGCGGGCTGTAGATCAGGTTTGGCCCACTTGAGTGAAGATG 900
1749 TTCAGTGAGAACCTGATCTNTCAGCCCAACACCCCACTTAAACGTTGGTGTAA 1808
901 TTCAGTGAGAACCTGATCTCTCAGCCCAACACCCCACTTAAACGTTGGTGTAA 1868
1809 NTGCCCTTCGGGTTGTCAGAAATNTAGCGAAGGCACTTTAAACGTTGGTGTAA 1928
961 CTGCCCTTCGGGTTGTCAGAAATNTAGCGAAGGCACTTTAAACGTTGGTGTAA 1020
1869 GAAGCTCTCAGGCGCCCAACGAGGTTGATCAGCTCAGTGGGAAAGAAATATAAT 1928
1021 GAAGCTCTCAGGCGCCCAACGAGGTTGATCAGCTCAGTGGGAAAGAAATATAAT 1080
1929 TTCCTTCAGGTTTAA 1946
1081 TTCCTTCAGGTTTAA 1098

RESULT 7
US-09-735-705-347
; Sequence 347, Application US/09735705
; Patent No. US2002052329A1

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong
APPLICANT: Fan, Liqun
APPLICANT: Kalos, Michael D.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Hosken, Nancy
APPLICANT: Fanger, Gary R.
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Henderson, Robert A.
APPLICANT: McNeill, Patricia D.
APPLICANT: Fanger, Neil

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.455C14

CURRENT APPLICATION NUMBER: US/09/735,705

CURRENT FILING DATE: 2000-12-12

NUMBER OF SEQ ID NOS: 419

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 347

LENGTH: 1740

TYPE: DNA

ORGANISM: Homo sapiens

US-09-735-705-347

Query Match 34.6%; Score 674.2; DB 9; Length 1740;

Best Local Similarity 56.7%; Pred. No. 6.6e-200;

Matches 980; Conservative 0; Mismatches 483; Indels 6; Gaps 1;

103 GGGTTAGGAGCGGAAAAATCAATCCGAATATTTCCACCCAGCTCCGATGGGAAGTAC 162
230 GGCMAAGGATTCGAAATCTTCAGATACGAATATCCCGCTCATTTACAGTGGGAGTGC 289
163 TGGACAGCCTCTGGCTCAGTATGGTACAGTAGAAGCTGTGAGCAGTGAACCCGAGA 222
290 TGGATAGTTTACCTAGTCCAGTATGGAGTGGGAGAGCTGTGAGCAGTGAACCTGACT 349
223 GTGAGACGAGAGTGGTGAATGTACCTTACCTATTTCCACCCGGGAGCAGACCAAGCATCA 282
350 CGGAAACTGCAAGTTGTAATGTAACCTTATTCAGTAAGGACCAAGCTAGACAGCCTAG 409
283 TGAAGCTGAATGGCCACACAGTTGAGAACCATGCTTGAAGGCTCTCTACATCCCGGATG 342
410 ACAACTGAATGGATTTTCAGTTAGAGAAATTCACCTTGAAGTAGCCTATATCCCTGATG 469
343 AGCAGATAGCAGGAGCCTGAGAAATGGGCCCGGAGGGGCTTTGGCTCTCGGGGTACG 402
470 AAACGGCGCCCAACAAAACCCCTTGAGCAGACCCCGAGGTCGCGGGGGCTTGGGCAGA 529
403 CCCGCCAGGGCTCAGCTGTGSCAGCGGGGGCCCGAGCCCAAGCAGCAGCAAGTGGACATCC 462
530 GGGGCTCTCTCAAGGACGAGGCTCTCAGGATCCGATCCGATCCAGAGCAAAACCATGTATTCG 589
463 CCCTTCGGCTCTCTGGTCCCGACCCAGTATGTGGGTGCCATTTATGGCAGAGGAGGGGCCA 522
590 CTCTGGCGCTCTGGTCTCCACCCAAATTTGTGGAGCCATCATAGGAAAGAGGTGCCA 649
523 CCATCCGCAACATCACAAACAGACCCAGTCCAGATAGAGTGCATAGGAGAGAGAGCG 582
650 CCATTCGGAACATCACCAACAGACCCAGTCTAAATCGATGTCCACCGTAAAGAAAATG 709
583 CAGGTGAGCTGAAAGACCATCAGTGTGCACTCCACCCCTGAGGGCTCTCTCCGCTT 642
710 CGGGGCTCTGAGAGTTCGATTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 769
643 GTAAGATCATCTTGAGAGATTATGCATAAAGAGGCTTAAGGACACCAAAACGGCTGACGAGG 702
770 GTAAGTCTATTCTGGAGATTATGCATAAAGAGGCTTAAGGACACCAAAACGGCTGACGAGG 829
703 TTCCCTCGAAGATCCTGGCCCATATAAATTTGTAGGGGCTCTCATTTGGCAAGGAGAGAC 762
830 TCCCTCTGAGATTCTTAGCTCATATAAATTTGTGGAGCTCTTATTTGGTAAAGAGGAA 889
763 GGAACCTGAAGAGGTAGACAGATACCGAGACAAATACCATCTCTCTCTCTCTCTCTCTCT 822
890 GAAATCTTAAAAAATTTGAGCAAGACACAGACATTAATACAGATATCTCCATTTGCAGG 949
823 ACCTTACCTTTTACAACTGAGAGGACCATCCTGTGAAAGGGGCGCATTCGAGAAATTTGT 882
950 AATTGACGCTGTATTAATCCAGAACGCACTTATACAGTTAAAGGCAATTTGAGACATGTG 1009
883 GCAGGCGCGAGCAGGAAATTAATGAAGAAAGTTTCGGAGGCGCTATGAGATGATGGCTG 942
1010 CCAGAGCTGAGGAGGAGATCATGAAGAAATCAGGAGTCTTATGAAATATGATTTGCTT 1069
943 CCATGA-----GCTCTACCTGATCCCTGGCTGACCTGCTGCTGAGTCTTTTTC 996
1070 CTATGAATCTTCAAGCACAATTTAATCTCTGGAATTAATGAACTGAAACGCTTTGGGTCTGTTCC 1129
997 CAGCTTTTATCCAGCGAGTCCCGCGCTCCCGAGCAGCGCTTACTTGGGGCTGCTCCCTATA 1056
1130 CACCACTTCAGGAGTGCACCTCCACCTCAGGGCCCCCTTCAGCCATGACTCTCTCCCT 1189
1057 GCTCTCTTTATGAGGCTCCCGAGCAGGAGATGGTGCAGTGTGTTTATCCCGCCCGAGGAG 1116
1190 ACCCGCAGTTTGAGCAATCAGAAAACGAGACTGTTTCATCTGTTTATCCAGCTCTATCAG 1249
1117 TGGGCGCATCATCGCAGAGAGGGGCGACGACATCAACAGCTCTCCCGGCTTGGCAGCG 1176
1250 TCGGTGCCATCATCGGCAAGAGGGGCGACGACATCAAGAGCTTTTCTCGTTTCTGCGAG 1309
1177 CTTCCATCAAGATTGACCAACCCGACCAACCTGACTCTCCAAAGTTCGTATGTTTATCATCA 1236


```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-327

Query Match      55.8%; Score 1085.4; DB 9; Length 1182;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1102; Conservative 0; Mismatches 6; Indels 6; Gaps 1;

QY 839 CCTGAGAGACCATCTCTGTGAAGGGGGCCATCGAATTTGTTCAGGCGCCGAGCAGGA 898
Db 18 CGCTGAGAGACCATCATCTGTGAAGGGGGCCATCGAATTTGTTCAGGCGCCGAGCAGGA 77

QY 899 AATAATGAAGAAATTCGGGAGGCTATGAGATGATGTGCTGCCATGAGC-----TC 952
Db 78 AATATGAGAGAGTTTCGGAGGGCTATGAGATGATGTGCTGCCATGAGCCTGCGAGTC 137

QY 953 TCACCTGATCCCTGGCTGACACCTGGCTGCTGTAGGTCTTTTCCAGCTTCATCCAGCGC 1012
Db 138 TCACCTGATCCCTGGCTGACACCTGGCTGCTGTAGGTCTTTTCCAGCTTCATCCAGCGC 197

QY 1013 AGTCCCGCGCTCCAGCAGCGTTACTGGGGTGTCTCCCTATAGCTTCCCTTATGCGAGC 1072
Db 198 AGTCCCGCGCTCCAGCAGCGTTACTGGGGTGTCTCCCTATAGCTTCCCTTATGCGAGC 257

QY 1073 TCCCGAGCAGGAGATGGTGAGGTGTTTATCCCGCCCGCAGCGAGTGGCGCCATCATCGG 1132
Db 258 TCCCGAGCAGGAGATGGTGAGGTGTTTATCCCGCCCGCAGCGAGTGGCGCCATCATCGG 317

QY 1133 CAAGAAGGGGACGACATCAAAACAGCTCTCCCGGTTTGGCAGCGCTCCATCAAGATTGC 1192
Db 318 CAAGAAGGGGACGACATCAAAACAGCTCTCCCGGTTTGGCAGCGCTCCATCAAGATTGC 377

QY 1193 ACCACCGGAAACCTGACTCCAAAGTTCGTATGTTTATCATCTGAGCGCCGAGCGC 1252
Db 378 ACCACCGGAAACCTGACTCCAAAGTTCGTATGTTTATCATCTGAGCGCCGAGCGC 437

QY 1253 CCAATTCAAGGCTCAGGAGAAATCTATGGCAAACCTCAAGGAGGAACTTCTTTGGTCC 1312
Db 438 CCAATTCAAGGCTCAGGAGAAATCTATGGCAAACCTCAAGGAGGAACTTCTTTGGTCC 497

QY 1313 CAAGGAGGAAGTGAAGCTGGAGCCACATACGTGTGCCAGCATCAGCAGCTGCGCGGT 1372
Db 498 CAAGGAGGAAGTGAAGCTGGAGCCACATACGTGTGCCAGCATCAGCAGCTGCGCGGT 557

QY 1373 CATTCGAAAGTGGAAACGGTGAACGTTGCGAATTTGACGGCAGCTGAGTGGT 1432
Db 558 CATTCGAAAGTGGAAACGGTGAACGTTGCGAATTTGACGGCAGCTGAGTGGT 617

QY 1433 AGTACCAAGAGACGACGCCCTGATGAGAACGACAGGTTCATCGTGAAATCATCGGACA 1492
Db 618 AGTACCAAGAGACGACGCCCTGATGAGAACGACAGGTTCATCGTGAAATCATCGGACA 677

QY 1493 TTTCTATGCGATCAGATGGCTCAACGGAAGATCCGAGACATCCTGGCCAGGTTAAGCA 1552
Db 678 TTTCTATGCGATCAGATGGCTCAACGGAAGATCCGAGACATCCTGGCCAGGTTAAGCA 737

QY 1553 GCAGCATCAGAGGGGACAGGTAAACAGCGCCAGGCAACGAGGAAATGACAGCCCTCC 1612
Db 738 GCAGCATCAGAGGGGACAGGTAAACAGCGCCAGGCAACGAGGAAATGACAGCCCTCC 797

QY 1613 CTGTCCCTTNGAGTCCAGGACAAACAGGGCGAGAAATCGAGAGTGTGCTCTCCCGGCGAG 1672
Db 798 CTGTCCCTTNGAGTCCAGGACAAACAGGGCGAGAAATCGAGAGTGTGCTCTCCCGGCGAG 857

QY 1673 GCCTGAGAAATGAGTGGGAATCCGGGACACNTGGCCGGGCTGTPAGATCAGGTTTGGCCAC 1732
Db 858 GCCTGAGAAATGAGTGGGAATCCGGGACACNTGGCCGGGCTGTPAGATCAGGTTTGGCCAC 917

QY 1733 TTGATTGAGAAAGATCTTCAGTGAAGAACCTTGATCTTTCAGCCGCCAACACCCACCCA 1792
Db 918 TTGATTGAGAAAGATCTTCAGTGAAGAACCTTGATCTTTCAGCCGCCAACACCCACCCA 977

QY 1793 ATTGGCCCAACACTGTNTGCCCTCGGGGTGTGAGAAATTTNTAGCGCAAGGCATTTTAA 1852
```

```
Db 978 ATTGGCCCAACACTGTCTGCCCTCGGGGTGTGAGAAATTTCTAGCGCAAGGCATTTTAA 1037
QY 1853 ACCTGATTGTTTAAAGAACTCTCCAGGCCCCACCAAGAGGGTGGATACACCTCAGTG 1912
Db 1038 ACCTGATTGTTTAAAGAACTCTCCAGGCCCCACCAAGAGGGTGGATACACCTCAGTG 1097
QY 1913 GGAAGAAAATAAAATTTCTTCCAGGTTTAAAA 1946
Db 1098 GGAAGAAAATAAAATTTCTTCCAGGTTTAAAA 1131

RESULT 6
US-09-764-864-745
; Sequence 745, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 745
; LENGTH: 1129
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-745

Query Match      55.0%; Score 1069.8; DB 9; Length 1129;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1084; Conservative 3; Mismatches 5; Indels 6; Gaps 1;

QY 855 ACTGTGAAGGGGGCCATCGAGATTTGTCGAGGCGGAGCAGGAATAATAGAAAGTT 914
Db 1 ACTGTGAAGGGGGCCATCGAGAAATTTGTCGAGGCGGAGCAGGAATAATAGAAAGTT 60

QY 915 CGGGAGGCTTATGAGAAATGATGTGGCTGCCATGAGC-----TCTCACCTGATCCCTGGC 968
Db 61 CGGGAGGCTTATGAGAAATGATGTGGCTGCCATGAGCCTGCAGTCTCCTGATCCCTGGC 120

QY 969 CTGAACCTGGCTGCTGTAGGTCTTTTCCAGTTCATCCAGCGAGTCCCGCGCTCCC 1028
Db 121 CTGAACCTGGCTGCTGTAGGTCTTTTCCAGTTCATCCAGCGAGTCCCGCGCTCCC 180

QY 1029 AGCAGGTTACTGGGGCTGCTCCCTATAGCTCTTTATGCGAGCTCCCGAGCAGGATG 1088
Db 181 AGCAGGTTACTGGGGCTGCTCCCTATAGCTCTTTATGCGAGCTCCCGAGCAGGATG 240

QY 1089 GTGCAAGTGTATATCCCGCCCGCAGGTGGGCGCATCATCGCAAGAGGGGCGAGCAC 1148
Db 241 GTGCAAGTGTATATCCCGCCCGCAGGTGGGCGCATCATCGCAAGAGGGGCGAGCAC 300

QY 1149 ATCAAAACAGCTCTCCGGTTTTCGAGGCTCCATCAAGATTGCAACCCGAAACACT 1208
Db 301 ATCAAAACAGCTCTCCGGTTTTCGAGGCTCCATCAAGATTGCAACCCGAAACACT 360

QY 1209 GACTCCAAAGTTCGTATGTTATCATCTGACCGCCAGAGGCCCAATTCAGGCTCAG 1268
Db 361 GACTCCAAAGTTCGTATGTTATCATCTGACCGCCAGAGGCCCAATTCAGGCTCAG 420

QY 1269 GGAAGAACTATGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGAGGAAAGTGAAG 1328
Db 421 GGAAGAACTATGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGAGGAAAGTGAAG 480

QY 1329 CTGGAGACCCACATACGTGTGCGCAGCATCAGCAGTGGCGGGTCAATTGGCAAAAGTGA 1388
Db 481 CTGGAGACCCACATACGTGTGCGCAGCATCAGCAGTGGCGGGTCAATTGGCAAAAGTGA 540

QY 1389 AAAACGGTGAACGCTTCAGAAATTTGACGCGAGCTGAGTGGTAGTACCAAGACACAG 1448
```

; SOFTWARE: PERL Program		
; SEQ ID NO 2082		
; LENGTH: 1738		
; TYPE: DNA		
; ORGANISM: Homo sapiens		
; NAME/KEY: misc feature		
; OTHER INFORMATION: Incyte ID No. US20040005559A1 457674.1		
US-10-062-674-2082		
Query Match		66.4%; Score 1291.4; DB 16; Length 1738;
Best Local Similarity		93.8%; Pred. No. 0;
Matches 1621; Conservative		0; Mismatches 61; Indels 46; Gaps 25;
QY	265	AGACCAGCGACCATCATGAAGCTGAATGCGCCACCACTGAGTGGAGAACCATCCCTCGAAGG 324
DB	1738	AGACGAGCGAAGCCATCATGAAGCTGAATGCGCCACCACTGAGTGGAGAACCATCCCTCGAAGG 1679
QY	325	TCTCTACATCCCGATGAGCAGATAGCACAGGACCTTGAGATGGGGCGCGAGGGGCT 384
DB	1678	TCTCTACATCCCGATGAGCAGATAGCACAGGACCTTGAGATGGGGCGCGAGGGGCT 1619
QY	385	TTGGCTCTCGGGTTCAGCCCGCCAGGGCTCACCTGTGGCAGCGGGGCCCCAGCCAGC 444
DB	1618	TTGGCTCTCGGGTTCAGCCCGCCAGGGCTCACCTGTGGCAGCGGGGCCCCAGCCAGC 1559
QY	445	AGCAGCAAGTGACATCCCTTCGGCTCCTGGTGGCCACCCAGTATGTGGGTGCCATTA 504
DB	1558	AGCAGCAAGTGACATCCCTTCGGCTCCTGGTGGCCACCCAGTATGTGGGTGCCATTA 1499
QY	505	TTGGCAAGGAGGGGCCACCATCCGCAACATCAAAAAAGACCCAGTCCCAAGATAGAGG 564
DB	1498	TTGGCAAGGAGGGGCCACCATCCGCAACATCAAAAAAGACCCAGTCCCAAGATAGAGG 1439
QY	565	TGCATAGAGAGAGAACCGAGCTGCAGCTGAAAGAGCCATCAGTGTCACTCCACCCCTG 624
DB	1438	TGCATAGAGAGAGAACCGAGCTGCAGCTGAAAGAGCCATCAGTGTCACTCCACCCCTG 1379
QY	625	AGGGTGCTCTCCG-CTTGTAAAGATGATCTTGGAGATTATGCATAAAGAGAGCTAAGAC 683
DB	1378	AGGGTGCTCTCTCCG-CTTGTAAAGATGATCTTGGAGATTATGCATAAAGAGAGCTAAGAC 1319
QY	694	ACCAAAACGGCTGACGAGGTTCCCTGAAAGTCCCTGGGCCATPAATACTTTGTAGGGGT 743
DB	1318	ACCAAAACGGCTGACGAGGTTCCCTGAAAGTCCCTGGGCCATPAATACTTTGTAGGGGT 1259
QY	744	CTC-ATTGGGGAAGAGACGGAACCTGAGNA-CGTAGACGAGATACCGAGACAAAA 801
DB	1258	CTCAATTGGCAAGGAAGACGGAACCTGAGGAAGGGTAGACCAATACCGAGACAAAA 1199
QY	802	TCACCATCTCTCTGTTGCAAGACCTTACCCCTTTACAA-CCCTGAGAGGACCATCACTGTG 860
DB	1198	TCACCATCTCTCTGTTGCAAGACCTTACCCCTTTACAA-CCCTGAGAGGACCATCACTGTG 1139
QY	861	AA-GGGGGCCATCGAGAAATTTGACAGG-CCGACGAGGAATAATGAAGAAAGTTGGG 918
DB	1138	RAAGGGGGCCATCGAGAAATTTGACAGG-CCGACGAGGAATAATGAAGAAAGTTGGG 1079
QY	919	AGGCTATGAGAAATGATGGTGTGCATGAGC-----TCTCACTGATCCCTGGCCTGA 972
DB	1078	AGGCTATGAGAAATGATGGTGTGCATGAGC-----TCTCACTGATCCCTGGCCTGA 1019
QY	973	ACCTGGCTCTGTAGTCTTTTCCAGCTTTCATCCAGCGCAGTCCCGCCGCTCCGACGA 1032
DB	1018	ACCTGGCTCTGTAGTCTTTTCCAGCTTTCATCCAGCGCAGTCCCGCCGCTCCGACGA 959
QY	1033	GGGTACTGGGGTGTCTCCTATAGTCTCCTTTATGCAAGCTCCCGAGGAGATGCTGC 1092
DB	958	GGGTACTGGGGTGTCTCCTATAGTCTCCTTTATGCAAGCTCCCGAGGAGATGCTGC 899
QY	1093	AGGTGTTATCCCGCCAGGAGTGGGGCCCATCATCGGCAAGAGGGGCGACATCA 1152
DB	898	AGGTGTTATCCCGCCAGGAGTGGGGCCCATCATCGGCAAGAGGGGCGACATCA 839

QY	1153	AAAGCT--CTCCGGTTTCCAGCGCTCC-ATCAAGATTGCACCCAGCAACACCTG 1209
DB	838	AAGCAGCTCTCTCCGGTTTCCAGCGCTCCAAATCAAGATTGCACCCAGCAACACCTG 779
QY	1210	ACTCCAAAGTTGTTATGTTATCATCTATGACCGCCAGAGG-CCCAATTCAAGGCTCAG 1268
DB	778	ACTCCAAAGTTGTTATGTTATCATCTATGACCGCCAGAGGCCCCAATTCAAGGCTCAG 719
QY	1269	GGAAAGATCTATGG-CAAATCAAGAGGAGAACTTTCTT-TGGTCCCAAGAGGAGTGA 1326
DB	718	GGAAAGATCTATGGCCAAATCTCAAGAGGAGAACTTTCTTGGGGTCCCAAGAGGAGTGA 659
QY	1327	AGCTGGAGACCCCATATAGTGTGCCAGCATCAGCAGCTGGCCGG--TCATTGGCAAGG 1384
DB	658	AGCTGGAGACCCCATATAGTGTGCCAGCATCAGCAGCTGGCCGGCTCATTGGCCAAAGG 599
QY	1385	TGGAAAAAC--GGTCAACGAGTTGCAGAAATTTGACGGCAGCTGAGTGGTAG--TACCA 1439
DB	598	TGGAAAAACGGGTGAA-CGAGTTGCAGAAATTTGACGGCAGCTGAGTGGTAGTACCAAG 539
QY	1440	AGAGACAGACCCCTGTATGAGAACGACAGGTCTATCGTGAATAATCATCGAC-ATTCTTA 1498
DB	538	AGACAGACCCCTGTATGAGAAACGACAGGTCTATCGTGAATAATCATCGCAATTTCTTA 479
QY	1499	TGCCAGTCAGATGCTCAACGGAA-GATCCGAGACATCTCGCCAGGTTAAGCAGCAGC 1557
DB	478	TGCCAGTCAGATGCTCAACGGAAAGGATCCGAGACATCTCGCCAGGTTAAGCAGCAGC 419
QY	1558	ATCAGAGGGACAGAGTAACCCAGGCCAGGACCGAGGAAGTGACAGCCCTC-CTGT 1616
DB	418	ATCAGAGGGACAGAGTAACCCAGGCCAGGACCGAGGAAGTGACAGCCCTCGCCTGT 359
QY	1617	CCCTTNGAGTCCAGGACAAACAAAGGGGAGAAATCAGAGTGTGTCTCTCCCGGAGGCT 1676
DB	358	CCCTTNGAGTCCAGGACAAACAAAGGGGAGAAATCAGAGTGTGTCTCTCCCGGAGGCT 299
QY	1677	GAGATGAGTGGGAATCCGGACAC-NTGGCCGGGCTGTAGATCAGGTTTCCCCAC-- 1732
DB	298	GAGATGAGTGGGAATCCGGGACACACTGGGCGGGCTGTAGATCAGGTTTCCCCACACT 239
QY	1733	--TTGATTTGAGAAAGATGTTCCAGTGAGAAACCTTGATCTTNTCAGCCCC--AAACA 1784
DB	238	GATTTGAGAAACGATGTTCCAGTGAGAAACCTTGATCTTNTCAGCCCCAACATACAC 179
QY	1785	CCCAACCAATTGG---CCCAACACTGTNTGCCCTCGGGGTGT-CAGAAATTTAGCGCA 1840
DB	178	CACACCAATTTGGCGCCCAACAACTGTCTGGGCTCGGGGTGTCCAGAAATCTAGCGCA 119
QY	1841	AG-GCACTTTTAAAGT-GGATTTGTTAAAGAGCTCTCCAGGCCCCCAAGAGGTTG 1898
DB	118	AGAGCACTTTTAAAGTGGGATTTGTTTAAAGAGCTCTCCAGGCCCCCAAGAGGTTG 59
QY	1899	ATCACACCTCAGTGGGAGAAAAATAAAATTTCTTTTCTTTCAGTTTAAAA 1946
DB	58	ATCACACCTCAGTGGGAGAAAAATAAAATTTCTTTTCTTTCAGTTTAAAA 11

RESULT 5
US-09-764-864-327
; Sequence 327, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 327
; LENGTH: 1182

Query Match		66.9%;	Score 1301;	DB 9;	Length 2224;
Best Local Similarity		83.5%;	Pred. No. 0;		
Matches 1556;		Conservative	0;	Mismatches	275;
				Indels	33;
				Gaps	6;
QY	108	AGGAGCCGGAAATTC	AAATTCGAAATTC	CCAAATTC	CCAAATTC
DB	365	AGGAGTCGGAAATAC	AGATCCGCAATAT	TCCACCT	CCGATCCGATGGAAATGCTAGAT
QY	168	AGCTGTCTGGCTCAG	TATGGTACAGTAGAGAA	CTGTGAGCAAGTGA	ACACCGAGAGTGAG
DB	425	AGCTGTCTGGCTCAG	TATGGTACAGTAGAGAA	CTGTGAGCAAGTGA	ACACCGAGAGTGAG
QY	228	ACGCGACTGTGTAAT	GCACCTATTC	CAACCGGAGCAGAC	CCAGGCAAGCATCATCAAG
DB	485	ACGCGGTGTCAAGT	CACTACTCTA	ACCGGAGCAGAC	CCAGGCAAGCATCATCAAG
QY	288	CTGAATGGCCACCA	GTGGAGAACCAT	GCCTGAAGGTCT	CTCTACATCCCGATGAGCAG
DB	545	CTAAATGGCCATCA	ACTGGAGAACCAT	GCCTGAAGGTCT	CTCTACATCCCGATGAGCAG
QY	348	ATAGCAGGACCTGA	GAATGGCGCGGAGGGG	CTTTGGCTCT	CGGGGTCAAGCCCGC
DB	605	ATAACCAAGGTCT	GAGAAATGGGCGTCT	GTGAGGGCTTT	TGGGTCTCGGGCGCAGCCCGG
QY	408	CAGGGCTCACCTG	TGGCAGCGGGGCC	CCAGCAAGCAGCAG	CAAGTGGACATCCCGCTT
DB	665	CAAGGTTCGCGGT	GCAGCAGGGGCT	CCAGCAAGCAGCAG	CAAGTGGACATCCCGCTT
QY	468	CGGCTCTGTGTGC	ACCCAGTATGTGGT	GGCCATTTAT	TGGCAAGGAGGGGCGCACCATC
DB	725	CGGCTCTGTGTGC	ACCCAGTATGTGGT	GGCCATTTAT	TGGCAAGGAGGGGCGCACCATC
QY	528	CGCAACATCAAAA	CAGACCCAGTCCA	AGATAGA	CGTGATAGGAGAGAACGAGGT
DB	785	CGAAACATCAAAA	CAGACCCAGTCCA	AGATAGA	CGTGATAGGAGAGAACGAGGT
QY	588	GCAGCTGAAAAAG	CCATCACTGTGCACT	CCACCCCTGAGGGCT	GTCTCTCCGCTGTGAG
DB	845	GCTGGGAGNAGCC	ATCAGCTGCAAT	CAACCCCTGAGGGCT	GTCTCTCCGCTGTGAG
QY	648	ATGATCTTGAGAT	TATGCAATAAAGAG	GGCTAAGGACACCA	AAACGGCTGACAGAGTTCCC
DB	905	ATGATCTTGAGAT	TATGCAATAAAGAG	GGCTAAGGACACCA	AAACGGCTGACAGAGTTCCC
QY	708	CTGAAGATCTGGC	CTATATACCTTT	TGAGGGCTCT	CATTTGGCAAGAGAGCGGAAC
DB	965	CTGAAGATCTGGC	CTATATACCTTT	TGAGGGCTCT	CATTTGGCAAGAGAGCGGAAC
QY	768	CTGAAGAGGTAG	AGCAAGATACCGGAC	CAAAAATCA	CCATCTCTCTGTTGCAAGACCTT
DB	1025	CTGAAGAGGTAG	AGCAAGATACCGGAC	CAAAAATCA	CCATCTCTCTGTTGCAAGACCTT
QY	828	ACCTTTACACCT	GAGAGGACCAT	CACTGTGAAGGGGCC	ATCGAGAAATTTGTCAGG
DB	1085	ACCTTTACACCT	GAGAGGACCAT	CACTGTGAAGGGGCC	ATCGAGAAATTTGTCAGG
QY	888	GCGGACGAGAAAT	ATAGAAAGTTCGGGAG	GCCTATGAGNA	TATGATGTGGTGCATG
DB	1145	GCGGACGAGAAAT	ATAGAAAGTTCGGGAG	GCCTATGAGNA	TATGATGTGGTGCATG
QY	948	AGC-----TCT	CACCTGATCCCTGG	CTGAACTGGCTGCT	GTAGTCTTTTCCAGCT
DB	1205	AGCTTGACGTCC	ACCTCATCTCCCT	GGCTTAACCTGGCTGCT	GTAGTCTTTTCCAGCT
QY	1002	TATTCAGGCGAGT	CCCGGCGCTCC	CAGAGCGTTACT	GGGCTGCTCCCTATAGCTCC
DB	1265	TATTCAGGCGAGT	CCCGGCGCTCC	CAGAGCGTTACT	GGGCTGCTCCCTATAGCTCC
QY	1062	TTTATCAGGCTCC	CCGAGGAGATGGT	GCAGGTGTTAT	CCCCCCCCCAGGAGTGGG
DB	1325	TTTATCAGGCTCC	CCGAGGAGATGGT	GCAGGTGTTAT	CCCCCCCCCAGGAGTGGG
QY	1122	GCCATCATCGG	CAAGAGGGGAGC	ACATCAAA	CAGCTCTCCCGGTTTGGCAGCGCTCC

RESULT 4

US-10-062-674-2082/c

; Sequence 2082, Application US/10062674

; Publication No. US2004000559A1

; GENERAL INFORMATION:

; APPLICANT: Loring, Jeanne F.; Kaeer, Matthew R.

; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS

; FILE REFERENCE: PA-0026-1 CIP

; CURRENT APPLICATION NUMBER: US/10/062,674

; CURRENT FILING DATE: 2002-01-30

; PRIOR APPLICATION NUMBER: US 09/625,102

; PRIOR FILING DATE: 2000-07-24

; NUMBER OF SEQ ID NOS: 2217

;; PRIOR APPLICATION NUMBER: US/09/061,709
;; PRIOR FILING DATE: 1998-04-17
;; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 5
; LENGTH: 1708
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-899-651-5

Query Match 85.8%; Score 1669.2; DB 9; Length 1708;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1676; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 268 CCAGGCAAGCCATCATGAAGCTGAATGGCCACCAAGTTGGAGAACATCGCCCTGAAGTCT 327
Db |||||
QY 30 CCCGGGAGCCATCATGAAGCTGAATGGCCACCAAGTTGGAGAACATCGCCCTGAAGTCT 89
Db |||||
QY 328 CCTACATCCCAGATGAGCAGATAGACAGGACCTGAGAAATGGGGCGCGAGGGGCTTTG 387
Db |||||
QY 90 CCTACATCCCAGATGAGCAGATAGACAGGACCTGAGAAATGGGGCGCGAGGGGCTTTG 149
Db |||||
QY 388 GCTCTCGGGGTACGCGCCCGCAGGGCTCACCTGTGGCAGCGGGGCCCGCCAGCAAGCAGC 447
Db |||||
QY 150 GCTCTCGGGGTACGCGCCCGCAGGGCTCACCTGTGGCAGCGGGGCCCGCCAGCAAGCAGC 209
Db |||||
QY 448 AGCAAGTGGACATCCCCCTTCGGCTCTCTGTGGTCCCGCCAGCTATGTGGGTGCCATTATTG 507
Db |||||
QY 210 AGCAAGTGGACATCCCCCTTCGGCTCTCTGTGGTCCCGCCAGCTATGTGGGTGCCATTATTG 269
Db |||||
QY 508 GCAAGGAGGGGGCCACCATCCGCAACATCAAAAACAGACCCAGTCCCAAGATAGACGTGC 567
Db |||||
QY 270 GCAAGGAGGGGGCCACCATCCGCAACATCAAAAACAGACCCAGTCCCAAGATAGACGTGC 329
Db |||||
QY 568 ATAGAGAGGAGAACGCGAGTGCAGTGAAGAGCCATCAGTGTGCATCCACCTCGAGG 627
Db |||||
QY 330 ATAGAGAGGAGAACGCGAGTGCAGTGAAGAGCCATCAGTGTGCATCCACCTCGAGG 389
Db |||||
QY 628 GCTGCTCTCGCTTGTGAAGATGATCTTGGAGATATGCAATAAGAGGCTAAGGACACCA 687
Db |||||
QY 390 GCTGCTCTCGCTTGTGAAGATGATCTTGGAGATATGCAATAAGAGGCTAAGGACACCA 449
Db |||||
QY 688 AAACGGCTGACGAGTTCCTCGTGAAGATCTCTGGGCCCAATAAATCTTTGAGGGGCTCTCA 747
Db |||||
QY 450 AAACGGCTGACGAGTTCCTCGTGAAGATCTCTGGGCCCAATAAATCTTTGAGGGGCTCTCA 509
Db |||||
QY 748 TTGGCAAGGAGGACGGAACCTGAAGAGGTTAGACGAGATACCGAGACAAAATCACCA 807
Db |||||
QY 510 TTGGCAAGGAGGACGGAACCTGAAGAGGTTAGACGAGATACCGAGACAAAATCACCA 569
Db |||||
QY 808 TCTCTCTGTGCAAGACCTTACCTTTTACAAACCTTGAGAGGACCATCATCTGTGAAGGGGG 867
Db |||||
QY 570 TCTCTCTGTGCAAGACCTTACCTTTTACAAACCTTGAGAGGACCATCATCTGTGAAGGGGG 629
Db |||||
QY 868 CCATCGAGAAATTTGTCAGGGCCGAGCGAAGAAATATGAAGAAAGTTGCGGAGGCGCTATG 927
Db |||||
QY 630 CCATCGAGAAATTTGTCAGGGCCGAGCGAAGAAATATGAAGAAAGTTGCGGAGGCGCTATG 689
Db |||||
QY 928 AGAATGATGTGCTGCGCATGAGCTCTACCTGATCCCTGGCCTGACCTGGCTGCTGTAG 987
Db |||||
QY 690 AGAATGATGTGCTGCGCATGAGCTCTACCTGATCCCTGGCCTGACCTGGCTGCTGTAG 749
Db |||||
QY 988 GTCTTTTCCAGCTTCATCCAGCGAGTCCCGCGGCTCCCGAGCAGCGTTACTGGGGCTG 1047
Db |||||
QY 750 GTCCTTTCCAGCTTCATCCAGCGAGTCCCGCGGCTCCCGAGCAGCGTTACTGGGGCTG 809
Db |||||
QY 1048 CTCCTTTATAGCTCTTTATGAGGCTCCCGAGCAGGAGATGTTGAGAGTGTATTATCCCG 1107
Db |||||
QY 810 CTCCTTTATAGCTCTTTATGAGGCTCCCGAGCAGGAGATGTTGAGAGTGTATTATCCCG 869
Db |||||
QY 1108 CCCAGGACGTGGGGCCATCATCCGCAAGAGGGGCGAGCAGATCAAAACAGCTCTCCCGGT 1167
Db |||||
QY 870 CCCAGGACGTGGGGCCATCATCCGCAAGAGGGGCGAGCAGATCAAAACAGCTCTCCCGGT 929
Db |||||

QY 1168 TTGCCAGGGCTCCATCAAGATTCACACCCGAAAACACCTGACTCCAAAGTTCGTATGG 1227
Db |||||
QY 930 TTGCCAGGGCTCCATCAAGATTCACACCCGAAAACACCTGACTCCAAAGTTCGTATGG 989
Db |||||
QY 1228 TTATCATCACTGGACCGCCAGAGGCCCAATTCAGAGGCTCAGGGAAGATCTATGGCAAC 1287
Db |||||
QY 990 TTATCATCACTGGACCGCCAGAGGCCCAATTCAGAGGCTCAGGGAAGATCTATGGCAAC 1049
Db |||||
QY 1288 TCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAGTGAAGTGAAGTGGAGACCCACATACGTG 1347
Db |||||
QY 1050 TCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAGTGAAGTGAAGTGGAGACCCACATACGTG 1109
Db |||||
QY 1348 TCCAGCATCAGCAGCTGGCGGCTCATTTGCAAGGTTGGAAGAACGCTGAACGAGTTGC 1407
Db |||||
QY 1110 TGCAGCATCAGCAGCTGGCGGCTCATTTGCAAGGTTGGAAGAACGCTGAACGAGTTGC 1169
Db |||||
QY 1408 AGAATTTGACCGGACAGCTGAGGTGTAGTACCAAGAGACCAAGCCCTGATGAGAACGACC 1467
Db |||||
QY 1170 AGAATTTGACCGGACAGCTGAGGTGTAGTACCAAGAGACCAAGCCCTGATGAGAACGACC 1229
Db |||||
QY 1468 AGGTATCGTGAATAATCATCGGACATTTCTATGCGAGTCAGATGGCTCAACGGAGATGCC 1527
Db |||||
QY 1230 AGGTATCGTGAATAATCATCGGACATTTCTATGCGAGTCAGATGGCTCAACGGAGATGCC 1289
Db |||||
QY 1528 GAGACATCTCGCCAGGTTAAGCAGCAGCATCAGAGGGACAGAGTAACAGGCGCCAGG 1587
Db |||||
QY 1290 GAGACATCTCGCCAGGTTAAGCAGCAGCATCAGAGGGACAGAGTAACAGGCGCCAGG 1349
Db |||||
QY 1588 CACGAGGAGAGTGACAGCCCTCTCCCTGTCCCTTNGAGTCCAGGACAAACAGCGGACGAA 1647
Db |||||
QY 1350 CACGAGGAGAGTGACAGCCCTCTCCCTGTCCCTTNGAGTCCAGGACAAACAGCGGACGAA 1409
Db |||||
QY 1648 ATCGAGAGTGTCTCTCCCGGAGGCTGAGATGAGTGGGAAATCCGGGACACNTGGGC 1707
Db |||||
QY 1410 ATCGAGAGTGTCTCTCCCGGAGGCTGAGATGAGTGGGAAATCCGGGACACNTGGGC 1469
Db |||||
QY 1708 CGGCTGTAGATCAGGTTTGGCCACTTCATTTAGAGAAAGATGTTCCAGTGAGGAAACCTCGA 1767
Db |||||
QY 1470 CGGCTGTAGATCAGGTTTGGCCACTTCATTTAGAGAAAGATGTTCCAGTGAGGAAACCTCGA 1529
Db |||||
QY 1768 TCTNTCAGCCCCAAACACCCCAATTTGGCCCAACACTGNTGCCCCCTCGGGGTGTCTAG 1827
Db |||||
QY 1530 TCTNTCAGCCCCAAACACCCCAATTTGGCCCAACACTGNTGCCCCCTCGGGGTGTCTAG 1589
Db |||||
QY 1828 AAATNTAGCGCAAGCGCACTTTTAAACGTGGATGTTTAAAGAGCTCTCCAGGCCCCAC 1887
Db |||||
QY 1590 AAATNTAGCGCAAGCGCACTTTTAAACGTGGATGTTTAAAGAGCTCTCCAGGCCCCAC 1649
Db |||||
QY 1888 CNAAGGGTGGATCAGCTCAGTCGGAAGAAAATAAAATTTCTTCAGGTTTAAAA 1946
Db |||||
QY 1650 CNAAGGGTGGATCAGCTCAGTCGGAAGAAAATAAAATTTCTTCAGGTTTAAAA 1708
Db |||||

RESULT 3

US-09-873-637-1
; Sequence 1, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-873-637-1

	D	b	61	 CCACCAGAGGCGGGGTGGAGGGCGAGTGTCTCAGCTTCCTCCGGGTTAGAGACCCGAAAA	120
	Q	y	121	TTCAAATCGAAATATTCCACCCCAGCTCCGATGGGAAGTAATGCAGACGCTGCTGGCTC	180
	D	b	121	TTCAAATCGAAATATTCCACCCCAGCTCCGATGGGAAGTAATGCAGACGCTGCTGGCTC	180
	Q	y	181	AGTATGGTACAGTAGAGAACTGTGAGCAAGTGAAACACCGAGAGTGCAGCGCAGTGGTGA	240
	D	b	181	AGTATGGTACAGTAGAGAACTGTGAGCAAGTGAAACACCGAGAGTGCAGCGCAGTGGTGA	240
	Q	y	241	ATGTCACCTATTCCAACCGGAGAGCAGACGAGCAAGCCATCATAGAAGTGAATGCCACC	300
	D	b	241	ATGTCACCTATTCCAACCGGAGAGCAGACGAGCAAGCCATCATAGAAGTGAATGCCACC	300
	Q	y	301	AGTTGAGAACCATGCCCTGAAGGTCCTCTCATCCCCGATGAGCAGATAGCACAGGAGC	360
	D	b	301	AGTTGAGAACAATGCCCTGAAGGTCCTCTCATCCCCGATGAGCAGATAGCACAGGAGC	360
	Q	y	361	CTGAAATGGGCGCCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGCAGGGCTCACCTG	420
	D	b	361	CTGAAATGGGCGCCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGCAGGGCTCACCTG	420
	Q	y	421	TGGCAGCGGGGCCCCACSCCAAGCAGCAGTAGTGGAATCCCCCTTCGGTCTCTGGTGC	480
	D	b	421	TGGCAGCGGGGCCCCACSCCAAGCAGCAGTAGTGGAATCCCCCTTCGGTCTCTGGTGC	480
	Q	y	481	CCACCAGTATGTGGTGCTCATTTATTTGGCAAGGAGGGGGCCACCATCCGCAACATCAAA	540
	D	b	481	CCACCAGTATGTGGTGCTCATTTATTTGGCAAGGAGGGGGCCACCATCCGCAACATCAAA	540
	Q	y	541	AACGACCCAGTCCAAGATAGACTGTCATAGGAAGGAGAACGCGAGTCAGCTGAAAAG	600
	D	b	541	AACGACCCAGTCCAAGATAGACTGTCATAGGAAGGAGAACGCGAGTCAGCTGAAAAG	600
	Q	y	601	CCATCAGTGTGACTCCACCCCTGAGGGCTGCTCTCCGCTTGTGAAGATGATCTTGGAGA	660
	D	b	601	CCATCAGTGTGACTCCACCCCTGAGGGCTGCTCTCCGCTTGTGAAGATGATCTTGGAGA	660
	Q	y	661	TTATGTCATAAAGAGCTAAGNACACAAACCGGCTGACGAGGTTCCTCTGAGATCCCTGG	720
	D	b	661	TTATGTCATAAAGAGCTAAGNACACAAACCGGCTGACGAGGTTCCTCTGAGATCCCTGG	720
	Q	y	721	CCCATAATPACTTTGTAGGGGCTCTCATTTGGCAAGGAGGAGCGAACTGAAAGAGGTAG	780
	D	b	721	CCCATAATPACTTTGTAGGGGCTCTCATTTGGCAAGGAGGAGCGAACTGAAAGAGGTAG	780
	Q	y	781	AGCAAGATCCGAGACAAATACCATCTCTCTGTTGCAAGACTTACCTTTTACAACC	840
	D	b	781	AGCAAGATCCGAGACAAATACCATCTCTCTGTTGCAAGACTTACCTTTTACAACC	840
	Q	y	841	CTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAAATGTTGTGAGGGCCGAGCAGAAA	900
	D	b	841	CTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAAATGTTGTGAGGGCCGAGCAGAAA	900
	Q	y	901	TATGAGAAAGTTGGAGGCCCTATGAGATGATGTGGCTGCCATGAGCTCTCACCTGA	960
	D	b	901	TATGAGAAAGTTGGAGGCCCTATGAGATGATGTGGCTGCCATGAGCTCTCACCTGA	960
	Q	y	961	TCCCTGGCCTGAACCTGGCTGTAGGTCTTTTCCAGCTTTCATCCAGGGCAGTCCCGC	1020
	D	b	961	TCCCTGGCCTGAACCTGGCTGTAGGTCTTTTCCAGCTTTCATCCAGGGCAGTCCCGC	1020
	Q	y	1021	CGCCTCCAGCAGGTTACTGGGGCTGCTCCCTATAGCTCTTTATCGAGGCTCCCGAGC	1080
	D	b	1021	CGCCTCCAGCAGGTTACTGGGGCTGCTCCCTATAGCTCTTTATCGAGGCTCCCGAGC	1080
	Q	y	1081	AGGAGATGGTGACAGGTGTTTTATCCCCCGCCAGGAGCAGTGGGGCCCATCATCGGCAAGG	1140
	D	b	1081	AGGAGATGGTGACAGGTGTTTTATCCCCCGCCAGGAGCAGTGGGGCCCATCATCGGCAAGG	1140
	Q	y	1141	GGCAGCACATCAACAGCTCTCCCGGTTTGGCCAGGCGCTCCATCAAGATTGCACCCCG	1200

Db	1141	GGCAGC	CATCAAA	CAGCTCTCCCGG	TTTGCACGGCCTCCATCAAGATTGCACCA	CCGG	1200	
QY	1201	AAA	CACCTGACTC	CAAA	AGTTCTGATGGTTATCATCACTGGACCGCCAGAGGCCCAATTC	A	1260	
Db	1201	AAA	CACCTGACTC	CAAA	AGTTCTGATGGTTATCATCACTGGACCGCCAGAGGCCCAATTC	A	1260	
QY	1261	AGG	CTCAGGAG	AAGATCTATGG	CAAACTCAAGAGGAGAACTTCTTTGGTCCCAAGAGG	1320		
Db	1261	AGG	CTCAGGAG	AAGATCTATGG	CAAACTCAAGAGGAGAACTTCTTTGGTCCCAAGAGG	1320		
QY	1321	AAG	TGAAGCTG	GAGACCCACA	TACGTCTGCCAGCATCAGCAGCTGGCCCGGTCATTGGCA	1380		
Db	1321	AAG	TGAAGCTG	GAGACCCACA	TACGTCTGCCAGCATCAGCAGCTGGCCCGGTCATTGGCA	1380		
QY	1381	AAG	GTGGA	AAAAACGGT	GAAACAGTTGACGGCAGCTGAGCTGGTAGTACCAA	1440		
Db	1381	AAG	GTGGA	AAAAACGGT	GAAACAGTTGACGGCAGCTGAGCTGGTAGTACCAA	1440		
QY	1441	GAG	ACACGACCC	CTGATGAG	AACGACACAGTCTCTGTA	AAAAATCATCGGACATTTCTATG	1500	
Db	1441	GAG	ACACGACCC	CTGATGAG	AACGACACAGTCTCTGTA	AAAAATCATCGGACATTTCTATG	1500	
QY	1501	CCAG	TCAGATGG	CTCAACG	GAAGATCCGACACATCCTGGCC	CAGGTTAAGCAGAGCATC	1560	
Db	1501	CCAG	TCAGATGG	CTCAACG	GAAGATCCGACACATCCTGGCC	CAGGTTAAGCAGAGCATC	1560	
QY	1561	AGA	AGGCACAG	AGTAA	CCAGGCCAGCAGGAGGAAGTGAC	CAGAGTGTCTCTCCCGGCAGGCTGAGA	1620	
Db	1561	AGA	AGGCACAG	AGTAA	CCAGGCCAGCAGGAGGAAGTGAC	CAGAGTGTCTCTCCCGGCAGGCTGAGA	1620	
QY	1621	TN	GAGTCCAGG	CAACAA	CCGGCAGAAATCGAGAGTGTCTCTCCCGGCAGGCTGAGA	1680		
Db	1621	TN	GAGTCCAGG	CAACAA	CCGGCAGAAATCGAGAGTGTCTCTCCCGGCAGGCTGAGA	1680		
QY	1681	AT	GAGTGG	GAATCGG	GACACNTGGCCGGCTGTAGATCAGGTTGCCAC	TTGATTGA	1740	
Db	1681	AT	GAGTGG	GAATCGG	GACACNTGGCCGGCTGTAGATCAGGTTGCCAC	TTGATTGA	1740	
QY	1741	GAA	AGATGTTCC	AGTGAG	GAACCTGATCTNTCAGCCCCAAACACCA	CCCAATTTGGCCC	1800	
Db	1741	GAA	AGATGTTCC	AGTGAG	GAACCTGATCTNTCAGCCCCAAACACCA	CCCAATTTGGCCC	1800	
QY	1801	AAC	ACTGNTG	CCCCCTCGGG	GTCTCAGAATTTAGCGCAAGGCATCTTTTAA	ACGTGGAT	1860	
Db	1801	AA	CAC	TGNTGCC	CCCTCGGGGTGTAGAAATTTAGCGCAAGGCATCTTTTAA	ACGTGGAT	1860	
QY	1861	TG	TTAAAGAG	CTCTC	AGGCCCCCAAGAGGGTGGATCACACCTCAGTGG	GAAGAAA	1920	
Db	1861	TG	TTAAAGAG	CTCTC	AGGCCCCCAAGAGGGTGGATCACACCTCAGTGG	GAAGAAA	1920	
QY	1921	AA	TAA	AAATTTCC	TTTCAGGTTTAAAA	1946		
Db	1921	AA	TAA	AAATTTCC	TTTCAGGTTTAAAA	1946		

```

RESULT 2
US-09-899-651-5
; Sequence 5, Application US/09899651
; Patent No. US20020111470A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tsung
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuch, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; TITLE OF INVENTION: Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/899,651
; CURRENT FILING DATE: 2001-07-06

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2004, 13:04:51 ; Search time 837.722 Seconds
(without alignments)
11355.361 Million cell updates/sec

Title: US-09-270-437D-7
Perfect score: 1946
Sequence: 1 gctgtacggagggtggg.....atttccttcaggttttaaaa 1946

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3216467 seqs, 2444149694 residues

Total number of hits satisfying chosen parameters: 6432934

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1941	99.7	1946	9	US-09-899-651-7
2	1669.2	85.8	1708	9	US-09-899-651-5
3	1301	66.9	2224	9	US-09-873-637-1
4	1291.4	66.4	1738	16	US-10-062-674-2082
5	1085.4	55.8	1182	9	US-09-764-864-327
6	1069.8	55.0	1129	9	US-09-764-864-745
7	674.2	34.6	1740	9	US-09-735-705-347
8	674.2	34.6	1740	9	US-09-850-716A-347
9	674.2	34.6	1740	9	US-09-897-778-347
10	674.2	34.6	1740	13	US-10-007-700-347
11	674.2	34.6	1740	15	US-10-117-982-347
12	674.2	34.6	1740	15	US-10-117-982-478
13	674.2	34.6	1740	16	US-10-313-986-347
14	674.2	34.6	1740	16	US-10-313-986-478

15	674.2	34.6	1743	9	US-09-897-778-447
16	674.2	34.6	1743	9	US-09-897-778-450
17	674.2	34.6	1743	13	US-10-007-700-447
18	674.2	34.6	1743	13	US-10-007-700-450
19	674.2	34.6	1743	15	US-10-117-982-447
20	674.2	34.6	1743	15	US-10-117-982-450
21	674.2	34.6	1743	16	US-10-313-986-447
22	674.2	34.6	1743	16	US-10-313-986-450
23	674.2	34.6	1764	9	US-09-850-716A-428
24	674.2	34.6	1764	9	US-09-897-778-428
25	674.2	34.6	1764	13	US-10-007-700-428
26	674.2	34.6	1764	15	US-10-117-982-428
27	674.2	34.6	1764	16	US-10-313-986-428
28	674.2	34.6	1799	16	US-10-313-986-485
29	674.2	34.6	34555	15	US-10-117-982-479
30	674.2	34.6	34555	16	US-10-313-986-479
31	672.6	34.6	4159	9	US-09-899-651-4
32	672.6	34.6	4181	9	US-09-735-705-175
33	672.6	34.6	4181	9	US-09-954-456-715
34	672.6	34.6	4181	9	US-09-850-716A-175
35	672.6	34.6	4181	9	US-09-897-778-175
36	672.6	34.6	4181	10	US-09-466-396A-175
37	672.6	34.6	4181	13	US-10-007-700-175
38	672.6	34.6	4181	15	US-10-117-982-175
39	672.6	34.6	4181	16	US-10-313-986-175
40	672.6	34.6	4434	13	US-10-116-802-145
41	664.6	34.2	1740	15	US-10-117-982-483
42	664.6	34.2	1740	16	US-10-313-986-483
43	585.4	30.1	3283	9	US-09-899-651-8
44	585.4	30.1	3667	15	US-10-097-340-146
45	585.4	30.1	3667	17	US-10-648-593-48

ALIGNMENTS

RESULT 1

US-09-899-651-7
; Sequence 7, Application US/09899651
; Patent No. US20020111470A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; TITLE OF INVENTION: Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/899,651
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 7
; LENGTH: 1946
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-899-651-7

Query Match 99.7%; Score 1941; DB 9; Length 1946;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1946; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GCTGTAGCGAGGGGCTGGGGGCTGCTCTGCCCCCTTCTTTCGCGGCTCGCGGCTCAGC	60
DB	1	GCTGTAGCGAGGGGCTGGGGGCTGCTCTGCCCCCTTCTTTCGCGGCTCGCGGCTCAGC	60
QY	61	CCACCACGAGGCCGGGGTGGGGGAGTGTCTCAGCTTCCCGGGTTAGGCGCGGAAA	120

This Page Blank (uspto)

Mon Jul 26 09:17:14 2004

Db 1683 GGTGAAGCAGCAGGAGCAGAATAACCCCTCAGGGAGTCGCCCTCAGAGCCGACCAAGTGAGG 1742

QY 1604 AGCCC 1608

Db 1743 CTCCC 1747

Search completed: July 24, 2004, 16:48:13
Job time : 143.138 secs

Db 1680 AAAAATTGTTAGTCTTAAGAGAGAGGTGAAATCTTGAAGCTCATATCAGAGTGCCATCCT 1739
QY 1357 CAGCAGCTGGCCGGTCAATGGCAAAAGGTGAAACCGGTGAACAGTTGAGAGTTTGA 1416
Db 1740 TTGCTGCTGGCAGAGTTATTGGAAAGAGGAGCAAAACGGTGAATGAACTTCAGAAATTGT 1799
QY 1417 CGCAGCTGAGGTGGTGTAGTACCAAGAGACCAAGACCCCTGATGAGAACCAACAGGTCAATCG 1476
Db 1800 CAAGTGCAGAGTTGTTGTTGCTCTGTCGACACACCTGATGAGAAATGACCAAGTGGTTG 1859
QY 1477 TGAATAATCATCGACATTTCTATGCCAGTCAAGTGTCAACGGAAGATCCGACATCC 1536
Db 1860 TCAAAATAAATGGTCACTTCTATGCTTCCAGGTTGCCAGAGAAAATTCAGGAATTC 1919
QY 1537 TGGCCCAAGGTTAAGCAGCAGCATCAGAAG 1565
Db 1920 TGAATCAGGTAAGCAGCAGCAACCAACAG 1948

RESULT 15

US-09-061-709-8
; Sequence 8, Application US/09061709B
; Patent No. 6297364
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
; FILE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/061,709B
; CURRENT FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 8
; LENGTH: 3283
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-061-709-8

Query Match 30.1%; Score 585.4; DB 3; Length 3283;
Best Local Similarity 64.1%; Pred.No. 1.4e-160;
Matches 964; Conservative 0; Mismatches 481; Indels 60; Gaps 3;

QY 104 GGTTAGAGCCGGAAATTCMAATCCGAAATATTCACCCAGCTCCGATGGGAAGTACT 163
Db 303 GCTAAGGAGCAGGAAATTCAGATTTCGAAATCCCTCTCCTCAGCTGCGAGGTGTT 362
QY 164 GGNACCTGCTGGCTCAGTATGGTACAGTAGAATCTGTGACAAAGTGAACACCGAGAG 223
Db 363 GGATGGACTTTTGGCTCAATATGGGACAGTGGAGAATGTGGAACAAGTCAACACAGACAC 422
QY 224 TGAGACGGCAGTGTGAATGTCACTTATTCACCCGGAGCAGACCGAGCCAGCCATCAT 283
Db 423 AGNAACCGCGTGTGCAAGTCAATATGCAACAGAGAGAAGAAAATAGCCATGGA 482
QY 284 GAAGTGAATGGCCACAGTTGGAGAACCATGCCCTGAAGGTCTCTACATCCCGCATGA 343
Db 483 GAAGCTAAGCGGCGATCAGTTTGAGAACTACTCTTCAAGATTTCTACATCCCGATGA 542
QY 344 GCAGATAGCACAGGGACCTGAGATGGGCGCGAGGGGGCTTGGCTCTCGGGGTGAGCC 403
Db 543 -----AGAGGTGAGCTCCCTTTGCGCCCTTCAGCGAGGCCAGCGTGGGAGCCACTC 593
QY 404 CCGCCAGGGCTCACTGTGGCAGCGGGGGCCCGCAGCCAGCAGCAAGTGGACATCCC 463
Db 594 TTCCCGGGAGCAGGCCACCGCCCTGGGGGCATCTCTCAGGGCCAGACGATGATTTCCC 653
QY 464 CTTGGGCTCTCTGGTCCCAACCCAGATATGTGGGTGCCATTATTGGCAAGGAGGGGCCAC 523

Db 654 GCTGCGAATCCTGTCCTCCACCCAGCTTTGTTGGTGCATCATCGGAAAGAGGCTTGAC 713
QY 524 CATCCGCAACATCACAAAACAGACCCAGTCCAAAGATAGAGTGCATAGGAAGGAGAACGC 583
Db 714 CATAAAGAAACATCACTTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTC 773
QY 584 AGGTGACGTGAAAGAACCCATCAGTGTGCACTCCACCCCTGAGGGGTGCTCTCCGCTTG 643
Db 774 TGGAGCTGCAGAGAAAGCCTGTCCACCATCCATGCCACCCAGAGGGGACTTCTGAAGCATG 833
QY 644 TAAGATGATCTTGGAGATATGATTAAGAGGCTTAAGACACCAAAACGGCTGACGAGGT 703
Db 834 CCGCATGATCTTGAATCATGAGAAAGAGGAGATGAGACCAAACTAGCCGAAGAGAT 893
QY 704 TCCCTGAAGATCCTGGCCCATATAATTTGTAGGCGCTCATTTGGCAAGGAAGACG 763
Db 894 TCCTCTGAATAATCTTGGCACAAATGGCTTGGTGGAGACTGATTGGAAAAGAGGCAG 953
QY 764 GAACCTGAAGAGGTGAGCAAGATACCGAGACAAAATCACCATCTCTCTGTTGCAAGA 823
Db 954 AAATTTGAAGAAAATTTGAACATGAAACAGGGACCAAGATAACAATCTCATCTTTGCAGGA 1013
QY 824 CTTTACCTTTACAACCTCAGAGGACCATCACTGTGAAGGGGCCCATCGAGAAATGTTG 883
Db 1014 TTTGAGCATATAACAACCGGAAAGAACCATCACTGTGAAGGGGCACAGTTGAGGCTGTGC 1073
QY 884 CAGGGCCGAGCAGAAAATAATGAAGAAAGTTCCGGAGGCGCTATGAGATGATGTGGCTGC 943
Db 1074 CAGTGTGATAGATAGATTTATGAAGAGCTGCTGAGGCTTTGAAAATGATATGTGGC 1133
QY 944 CATGAGCTCTACCTGATCCCTGGCTGAACTGTGGCTGTGTAGGTCTTTTCCAGCTTC 1003
Db 1134 TGTTAACACCCACTCCGGATACCTTCTCCAGCCTG-----TAC 1170
QY 1004 ATCCAGCGAGTCCCGCGCTCCACAGCAGCGTTACTTGGGGCTGCTCCCTATAGCTCCTT 1063
Db 1171 CCCATCAACAGTTTGGCCCGTTTCCCGCATCACTCTTA----- 1211
QY 1064 TATGAGGCTCCCGAGCAGAGATGTCAGAGTGTATATCCCGCCGAGCAGTGGCGC 1123
Db 1212 -----TCCAGAGCAGAGATTTGTGAATCTTTCATCCCAACCCAGGCTGTGGCGC 1262
QY 1124 CATCATCGCAAGAGGGGAGCACAATCAACAGCTCTCCGGTTTGCAGAGCCCTCCAT 1183
Db 1263 CATCATCGGGAAGAGGGGGGCACATCAACAGCTGGCGAGATTCGCGGAGCTCTAT 1322
QY 1184 CAAGATTGCACCAACCGAAACACCTGACTCCAAAGTTCTGATGTTTATCATCTGAC 1243
Db 1323 CAAGATTGCCCTGCGGAAGGCCAGAGCTCAGCGAAAGATGGTTCATCATCACCGGGC 1382
QY 1244 GCCAGAGGCCAAATTCAGGCTCAGGGAAGATCTATGCAAACTCAAGAGAGAGAACTT 1303
Db 1383 ACGGAAGCCAGTTCAAGGCCAGGAGGATCTTTGGGAACCTGAAAGAGGAAACTT 1442
QY 1304 CTTTGTCCCAAGAGAGAGTGAAGCTGAGACCCACATACGTGTGCCAGCATCAGCAGC 1363
Db 1443 CTTTAAACCCCAAGAAAGATGAAGCTGGAAGCGCATATCAGAGTGCCCTCTTCACAGC 1502
QY 1364 TGGCCGGGTCTATTGGCAAGGTGGAACCGGTGAACGTTTGCAGATTTGACCGCAGC 1423
Db 1503 TGGCCGGGTGATTGGCAAGGTGGCAAGCCCGTGAACCGAATTCGAGAACTTAAACAGTGC 1562
QY 1424 TGAGTGGTGTAGTACCAAGAGACAGACCCCTGATGAGAACCGACAGGTCATCGTGAAT 1483
Db 1563 AGAAGTCATCGTGCCTCGTGACCAACCGCCAGATGAATAGGAGAGTGCATCGTCAAT 1622
QY 1484 CATCGCAATTTCTATGCCAGTCAAGTGGCTCAACGGAAGATCCGAGACATCTCGGCCCA 1543
Db 1623 TATCGGGCACTTCTTTGTGTAGCCAGCTGCACAGCGCAAGATCAGGGAAATTTGTAACA 1682
QY 1544 GGTTAAGCAGCAGCATCAGAGGGGACAGATTAACAGGCCCCAGGCCAGGAGAGTGAAC 1603

LOCATION: (3502)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3506)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3520)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3538)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3549)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3646)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3940)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3968)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3974)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (4036)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (4056)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (4062)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (4080)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (4088)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (4115)
OTHER INFORMATION: n=A,T,C or G
US-09-606-421B-175

Query Match 34.6%; Score 672.6; DB 4; Length 4181;
Best Local Similarity 66.6%; Pred. No. 5.4e-186;
Matches 979; Conservative 0; Mismatches 484; Indels 6; Gaps 1;
QY 103 GGGTTAGGAGCGGAAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTAC 162
DB 480 GGCAAAGGATTCGGAACCTTCAGATACGAAATATCCCGCTCATTTACAGTGGGAGGTGC 539
QY 163 TGGACAGCTGCTGGCTCAGTATGGTACAGTACGAACTGTGAGCAAGTGAACACCGAGA 222
DB 540 TGGATAGTTTACTAGTCCAGTATGGTGGAGAGCTGTGAGCAAGTGAACACTGACT 599
QY 223 GTGAGAGCGGAGTGGTGAATGTCACTATTTCCAAACCGGGAGAGACCGAGCAAGCCATCA 282

DB CGGAACTGCAGTTGTAAATGTAACCTATTCAGTAAGGACCAAGCTAGACAAGCACTAG 659
QY TGAAGCTGAATGGCCACAGTTGGAGAACATGCCCTGAAGGTCTCTTACATCCCGCATG 342
DB ACAAAGTGAATGGATTTCACTTGAAGAAATTTCACTTGAAGTACCTATATCCCTGATG 719
QY AGCAGATAGCACAGGACCTGAGAAATGGGCGCCGAGGGGCTTTGGCTCTCGGGCTCAGC 402
DB AATGGCGCCCGACGAAACCCCTTGCAGACAGCCCGGAGTCCGCGGGGCTTTGGGAGA 779
QY CCGCCAGGGCTCACCTGTGGCAGCGGGGCGCCAGCCAGCAAGCAGCAGTGGACATCC 462
DB GGGGCTCTCAAGCAGGGGTCTCCAGATCCGTATCCAAGCAGAAACCATGTGATTGC 839
QY CCTTTCGGCTCTCGTGGTCCACCCAGTATGTGGTGGCCATTTATGGCHAGAGGGGCCA 522
DB CTCTGGCCTGCTGGTTCCCAACCAATTTGTGGAGCCATCATAGGAAAGAGGTGCCA 899
QY CCATCCGCAACATCACAAACAGACCCAGTCCAGATAGAGTGCATAGGAGGAGAAACG 582
DB CCATTCGGAACATCACAAACAGACCCAGTCCAGATAGAGTGCATAGGAGGAGAAACG 959
QY CAGGTGCAGCTGAAAGAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCTCGCTT 642
DB CGGGGCTGCTGAGAGTCCGATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1019
QY GTAAGATGATCTTGGAGATTTGATTAAGAGGCTAAGGACACCAAAACGGCTGACGAGG 702
DB GTAAAGTCTATTCTGGAGATTTGATTAAGAGGCTAAGGACACCAAAACGGCTGACGAGG 1079
QY TTCCCTCGAAGATCCTGGCCCATTAATCTTTGTAGGGCTCTCTCTCTCTCTCTCTCTCT 762
DB TCCCTTGAAGATTTAGCTCATATAAATCTTTGTGGAGCTTTATTTGGTAAGAGGAA 1139
QY GGAACCTGAAGAGGTAGAGCAAGATACCGAGCAAAATACCAATCTCTCTCTCTCTCTCT 822
DB GAAATCTTAAAGAAATTTAGCAAGACACAGACACTAAATACGATATCTCCATTGCAGG 1199
QY ACCTTACCTTTTAAACCTTGAGAGGACCATCATCTGTGAAGGGGGCCATCGAGATTTGT 882
DB AATTGACGCTGTATAATCCAGAACGCACTATTACAGTTTAAAGGCAATGTGAGACATGTG 1259
QY GCAGGCGGAGCAGGAAATTAAGAAAGTTTCGGGAGGCGCTATGAGATGATGTGGCTG 942
DB CCAAAGCTGAGGAGAGATCATGAAGAAATCAGGAGTCTTATGAATGATATTTGCTT 1319
QY CCATGA-----GCTCTCACCTGATCCCTGGCTGAAACCTGGCTGTGTAGTCTTTTCC 996
DB CTATGAATCTTCAAGCACTTTAATCTCTGGATTAATCTGAACGCTTTGGGTCTGTCTCC 1379
QY CAGCTTTCATCCAGCGAGTCCCGCGCTCCAGCAGCGTTACTTGGGGCTGCTCCCTATA 1056
DB CACCCACTTTCAGGGATGCCACCTCCACCTCAGGGCGCCCTTCAGCCATGACTCTCTCCCT 1439
QY GCTCTTTATGAGGCTCCGAGCAGGAGATGGTGCAGTGTATTATCCCGCCCGCAGGAG 1116
DB ACCCGCAGTTTGAGCAATCAGAAACGAGAGCTTTCATCAGTTTATCCAGCTCTATCAG 1499
QY TGGGGCCATCATCGCAAGAGAGGGGAGCAGCATCAAAAGCTCTCTCCCGTTTGGCAGCG 1176
DB TCGGTGCCATCATCGCAAGAGGGGAGCAGCATCAAGCAGCTTTCTCGCTTTGGTGGAG 1559
QY CTCTCATCAAGATTGCACCCCGAAACACCTGACTCCAAAGTTCGTAAGTTTATCATCA 1236
DB CTTCAATTAAGATTGTCCAGCGGAGACCAAGATGCTAAAGTGAAGTGGTGAATATCA 1619
QY CTGGACCGCAGGCGCCCAATTAAGGCTCAGGGAAGATCTATGCAAACTCAAGAGG 1296
DB CTGGACCGCAGGCGCTCAGTTCAGGCTCAGGGAAGAAATTTATGAAGAAATTAAGAG 1679
QY AGAATCTCTTTGGTCCCAAGGAGAGTGAAGCTGGAGAGCCCAATCATCGTGTCCAGCAT 1356

```

; LOCATION: (4080)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (4088)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (4115)
; OTHER INFORMATION: n=A, T, C or G
US-09-542-615A-175

```

Query Match	34.6%	Score 672.6;	DB 4;	Length 4181;
Best Local Similarity	66.6%	Pred. NO. 5.4e-186;		
Matches 979; Conservative	0;	Mismatches 484;	Indels 6;	Gaps 1;

103	QY	GGGTTAGGACCGGAAATTCAAATCCGAATATATCCACCCAGCTCCGATCGGAAGTAC	162
480	Db	GGCAAGAGATTCCGAAACTTCAGATACGAATATCCCGCCTCAITTTACAGTGGGAGTGC	539
163	QY	TGCACAGCCTCTGGCTCAGTATCGGTACAGTAGAGAACTGTGACMAAGTGAACCCAGA	222
540	Db	TGGATAGTTTACTAGTCCAGTATCGAGTGGTGGAGAGCTGTGACCAAGTGAACACTGACT	599
223	QY	GTGAGACGGCAGTGGTGAATGTCCACTATTCCAAACGGGAGCAGACAGGCAAGCCATCA	282
600	Db	CGGAAACTGCAGTTGTAAATGTAACCTATTCCAGTAAGGACCAAGCTAGACAGCACTAG	659
283	QY	TGAAGCTGAATGCCACAGTTCGGAGAACCATGCCCTGAAGTCTCTCATATCCCCGATG	342
660	Db	ACAACTGAATGGAATTCAGTTAGAGAAATTCACCTTGAAGTAGGCTATATCCCTGATG	719
343	QY	AGCAGATAGCACAGGGACCTTGAGAAATGGGCGCGAGGGGGCTTTGGCTCTCGGGGTGAGC	402
720	Db	AAATGGCGCCACGACAAACCCCTTGACAGCAGCCCGAGTTCGCCGGGCTTGGGCAGA	779
403	QY	CCGCGCAGGGCTCACTGTGGCAGCGGGGGCCCCAGCGAGCGAGTAGTGGACATCC	462
780	Db	GGGGCTCTCAAGGACAGGGGTCTCCAGGATCCGTATCCAAAGCAGAAACCATGTGATTTGC	839
463	QY	CCCTTCGGTCTCTGGTGGCCACCAGTAGTGTGGTGCCATTATTGGCAAGGAGGGGGCCA	522
840	Db	CTCTGCGCTGCTGGTTCCACCCAAATTGTTGGAGCCATCATAGGAAGAAGGTGCCA	899
523	QY	CCATCCGACATCAACAACAGACCCAGTCCAAAGATAGACGTGCATAGGAAGGAGAACG	582
900	Db	CCATTCCGAAACATCACCAACAGACCCAGTCTAAAATCGATGTGCCCGTAAAGAAAATG	959
583	QY	CAGGTCCAGCTGAAAAAGCCATCAGTGTGCACCTCCACCTGAGGGCTGCTCCTCGCTT	642
960	Db	CGSGGGCTGCTGAGAAGTCGATTACTATCCTCTCTACTCTCTGAAGGACCTCTGCCGCTT	1019
643	QY	GTAAGATGATCTTGGAGATTATGCTAAGAGGCTAAGGACACCAAAACGGCTGACGAGG	702
1020	Db	GTAAGTCTATTCTGGAGATTATGCTAAGGAAGCTCAAGATATAAAAATTCACAGAAGAGA	1079
703	QY	TTCCCTCGAAGTCTCTGGCCCATATAACTTTGTAGGGGCTCTCATTTGCCAAGGAGAC	762
1080	Db	TCCTCTTGAAGATTTAGTCTCATATACCTTTGTTGGAGCTCTATTGGTAAAGAGAGAA	1139
763	QY	GGAACTGAAGAGGTAGACGAGATACCGAGACAAAAATCACCATCTCTCTGTTCCAAG	822
1140	Db	GAAATCTTTAAAAAATTGAGCAAGACACAGACATAAAATCAGATATCTCCATTGCAGG	1199
823	QY	ACCTTACCTTTACACCTGTGAGAGACCATCACTGTGAAGGGGGCCATCGAGAAATGTT	882
1200	Db	AAATTGACGCTGTATATCCAGAACGCATATTACAGTTAAAGGCATGTTGACATGTG	1259
883	QY	GCAGGGCCGAGCAGGAAATAATGAAGAAAGTTTCGGGAGGCCCTATGAGATGATGTGGCTG	942
1260	Db	CCAAAGCTGAGGAGGAGATCATGAAGAAATCAGGAGTCTTATGAAATGATTTGCTT	1319
943	QY	CCATGA-----GCTCTACCTGATCCCTGGGCTGAACTGGCTGCTGTAGTCTTTTCC	996
1320	Db	CTATGAATCTCAAGCACATTTAAATCTCTGATTAATCTGAACCGCTTGGGTCTGTTCC	1379

Db 720 AATATGCGCCGACGACAAACCCCTTGCAGCAGCCCGAGGTGCGCGGGGGCTGGGGCAGA 779
QY 403 CCGCGCAGGCTCACCTGTGGCAGCGGGGCCCGCAGCGAAGCAGCAGCAAGTGCACATCC 462
Db 780 GGGGCTCTCAAGCGAGGGTCTCCAGGATCCGATCCAGCAGCAACCATGTGATTGC 839
QY 463 CCCTTCGGCTCTCGGTGCCACCCAGTATGCGGTGCTATTTGGCAAGGAGGGGCCA 522
Db 840 CTCTGCGCTGCTGTTCCCAACCAATTTGTGGAGCCATCATAGGAAAGAGGTGCCA 899
QY 523 CCATCCGCAATCATCAAAAACAGACCAGTCCAAAGATAGAGTGCATAGGAAGAGAACG 582
Db 900 CCATTCGGAACATCACCAAAACAGACCAGTCTAAATTCGATGTCACCGTAAAGAAATG 959
QY 583 CAGTGCAGCTGAAAGGCAATCAGTGTCACTCCAGCCCTGAGGGTGTCTCCCGCTT 642
Db 960 CGGGGCTGCTGAGAGTCAATCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1019
QY 643 GTAAGATGATCTTGGAGATTATGATAAAGAGGCTAAGGACACCAAAACGGCTCAGCAGG 702
Db 1020 GTAAGTCTATTCGAGATTATGCATAAAGGAAGCTCAAGATATATAAATTCACAGAAGA 1079
QY 703 TTCCCTGAGATCTCTGGCCATATTAATCTTTAGGCGTCTCATTTGGCAAGGAGGAC 762
Db 1080 TCCCTTGAAGATTCTAGCTCATATAAATCTTTGTGGACGCTCTATTGGTAAAGAGAA 1139
QY 763 GGAACCTGAAGGATGAGCAAGATACCGAGACAAAATCACCATCTCTCTGTTGCAAG 822
Db 1140 GAAATCTTAAAAAATTGAGCAAGACACAGACATTAATACGATATCTCAATTCGAGG 1199
QY 823 ACCTTACCTTTTACAACCTCGAGAGGACCATCATCTGTGAAGGGGGCCATCGAGAAATGTT 882
Db 1200 AATTGACGCTGATATCCAGAACGCACTATTACAGTTAAAGGCAATGTTGAGACATGTG 1259
QY 883 CGAGGCGCAGAGAAATATGAAGAAGTTCCGGAGCCTATCAGAAATGATGCGGTG 942
Db 1260 CCAAAGCTGAGGAGAGATCATGAAGAAAATCAGGGAGTCTTATGAAAATGATATGCTT 1319
QY 943 CCATCA-----GCTCTCACCTGATCCCTGGCTCAACCTGGCTGCTGTAGTCTTTTCC 996
Db 1320 CTATGAATCTTCAAGCACAATTAATCTCTGATTAATCTGAACTTGGCTGTGTTCC 1379
QY 997 CAGCTTTCATCCAGCGAGTCCCGCGCTCCAGCAGCGTTTACTCGGGGTGCTCCCTATA 1056
Db 1380 CACCACTTCAGGATGCCACCTCCACTCAGGCGCCCTTCAGCCATGACTCTCCCT 1439
QY 1057 GCTCTTTATGAGGCTCCGAGCAGGATGCTGAGGTGTTATCCCGCCAGGAG 1116
Db 1440 ACCCGCAGTTTGAAGCAATCAGAAACGAGACTGTTATCAGTGTATCCCGAGCTCTATCAG 1499
QY 1117 TGGGCGCATCATCGCAAGAGGGGCGAGCAGCATCAACAGCTCTCCCGTTTGGCAGG 1176
Db 1500 TCGGTGCCATCATCGCAAGCAGGGCCAGCAGCATCAAGCAGCTTCTCGCTTGTGGAG 1559
QY 1177 CTTCCATCAGATTGCAACCCAGCAACCTGACTCCAAAGTGTGATGGTTATCATCA 1236
Db 1560 CTTCAATTAAGATTGCTCCAGCGGAAGCACCAGATGCTAAAGTGAGGATGGTGAATTATCA 1619
QY 1237 CTGGAACCGCAGAGGCCAATTCAGGCTCAGGAGAGATCTATGCCAACTCAAGGAGG 1296
Db 1620 CTGGACCAACAGAGGCTCAGTTCAAGGCTCAGGGAGAAATTTATGGAATAAAGAG 1679
QY 1297 AGAATCTCTTGTCTCCAGGAGGAGTGAAGCTGGAGACCCACATACGTTGCCAGCAT 1356
Db 1680 AAAATCTTGTAGTCTTAAGAGAGGTGAACCTTGAAGCTCATATCAGAGTGCATCCT 1739
QY 1357 CAGCAGCTGCGCGGTTCATTGGCAAAAGGTGGAAAAACCGTGAACGAGTTGCAGAAATTGA 1416
Db 1740 TTGCTGCTGCGCAGAGTTATTGGAAAAAGGAGGCAACCGTGAATGAATTCAGAAATTGT 1799
QY 1417 CGGCACTGAGGTGTTAGTACCAAGAGCAGACCCCTGATGAGAACGACCCAGTCACTG 1476
Db 1800 CAAGTCAGAAGTTGTTGTTCCCTCGTGACCAAGACACCTGATGAGATGACCAAGTGTTG 1859

QY 1477 TGAATATATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCC 1536
Db 1860 TCAAAATAACTGGTCACTTCTATGCTTGCAGGTTGCCAGAGAAAATTCAGGAATTC 1919
QY 1537 TGGCCCAAGGTTAAGCAGCAGCATCAGAAG 1565
Db 1920 TGACTAGGTAAGCAGCACCACCAACAG 1948

RESULT 13
US-09-542-615A-175
; Sequence 175, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fangez, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 175
; LENGTH: 4181
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3347)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (3502)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (3506)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (3520)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (3538)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (3549)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (3646)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (3940)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (3968)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (3974)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (4036)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (4056)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (4062)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure

```
QY 163 TGGACAGCGCTGCTGGCTCAGTATGTTACAGTAGAGAACTGTGAGCAAGTGAACACCGAGA 222
Db 540 TGGATAGTTTACTAGTCCAGTATGAGTGTGGAGAGCTGTGACAAAGTGAACACTGACT 599
QY 223 GTGAGACGGAGTGGTGAATGTCTACCTATTCCAAACCGGAGAGCAGACACAGGCAAGCCATCA 282
Db 600 CGGAAACTGCGAGTGTGAATGTAACTTATCCAGTAAGGACCAAGCTAGACAGCACTAG 659
QY 283 TGAAGCTGAATGGCCACAGTGTGAGAAACCATGCCCTGAAGTCTCTCTACATCCCGCATG 342
Db 660 ACAAACCTGAATGGATTTTCAGTTAGAGAAATTTCACTTTGAAAGTAGCCCTATATCCCTCATG 719
QY 343 AGCAGATAGCACAGGAGACCTGAGAAATGGGCGCCGAGGGGCTTTGGCTCTCGGGGTACG 402
Db 720 AATGGCGCGCCAGCAAAACCCCTTGCAGCAGCCCCAGGTGCGGGGGCTTTGGGCGA 779
QY 403 CCGCGCAGGGCTCACTGTGTGGCAGCGGGGCCCCAGCCAAAGCAGCAGCAAGTGGACATCC 462
Db 780 GGGGCTCTCTCAGGCGAGGGGTCTCCAGGATCCGATCCAAGCAGAAACCATGTGATTTC 839
QY 463 CCTTCGGCTCCTGGTGCACCCAGTATGTGGTGCATATTTGGCAAGGAGGGGCCA 522
Db 840 CTCTGGCTCTGCTGGTGTCCACCCAAATTTGTGGAGCCATATAGGAAAGAGGTGCCA 899
QY 523 CCATCCGCAACATCACAAAACAGACCCAGTCCAAAGATAGACGTGCATAGGAAGAGAAAG 582
Db 900 CCATTCCGGAACATCACAAACAGACCCAGTCTAAATCGATGTCCACCGTAAGAAATG 959
QY 583 CAGGTGAGCTGMAAAGCCATCAGTGTGCACCTCCACCCCTGAGGGTGTCTCTCGCTT 642
Db 960 CGGGGGTGTGAGAAATGCTACTATCTCTCTACTCTGAAAGCACCTCTGCGCTT 1019
QY 643 GTAAGATGATCTTGGAGATTCATTAAGAGGCTAAGGACACCAAAACGGCTGACAGG 702
Db 1020 GTAAGTCTATCTGGAGATTCATTAAGAAAGCTCAAGATATATAAATTCACAGAGAGA 1079
QY 703 TTCCCTCAGAGATCTCTGGCCATAATAACTTTGTAGGGCTCTCATTTGGCAAGGAAGGAC 762
Db 1080 TCCCTTTGAAGATTTTAGTCTAATAACTTTGTGGAGTCTTATTTGTAAGAGGAA 1139
QY 763 GGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAAAATCAACATCTCTCTGTTCAAG 822
Db 1140 GAAATCTTAAAGAAATTTGAGCAAGACACACACTAAATCAAGATATCTCCATTGCGG 1199
QY 823 ACCTTACCCTTTACACCTGTGAGAGCACCATCACTGTGAGGGGCCATCGAGAAATGTT 882
Db 1200 AATTGACGTGTATATCCAGAACGCACTATTACAGTTAAAGCAATGTGAGACATGTG 1259
QY 883 GCAGGCGCGAGCAGGAAATAATGAAGAAAGTTCCGGAGGCCTATGAGAAATGATGTGCTG 942
Db 1260 CCAAGCTGAGGAGGAGATCATGAAGAAATCAGGGAGTCTTATGAAGATGATTTGCTT 1319
QY 943 CCATGA-----GCTCTACCTGATCCCTGGCTGAACTGGCTGTGTAGGCTTTTCC 996
Db 1320 CTATGAATCTTCAAGCACATTTAAATCTCTGGATTAATCTGAACGCTTTGGGCTCTGTTCC 1379
QY 997 CAGCTTCAATCCAGCGCAGTCCCGCCCTCCAGCAGCGCTTACTGGGCTGTCTCCATTA 1056
Db 1380 CACCCACTTCAGGATGCCACCTCCCACTCAGGGCCCCCTTCAGCCATGACTCTCCCT 1439
QY 1057 GCTCTCTTATCAGGCTCCCGAGCAGAGATGGTGCAGGTGTTATCTCCCGCCAGGAG 1116
Db 1440 ACCCGCAGTTTGAGCAATCAGAAACGAGACTGTTTCATCAGTTTATCCAGCTCTATCAG 1499
QY 1117 TGGGCGCATCATCGCAAGAGGGCGCAGCATCAAAACAGTCTCTCCGGTTTGCAGCG 1176
Db 1500 TCGGTGCCATCATCGCAAGCAGGGCGCAGCATCAAGCAGCTTCTCGCTTTGCTGGAG 1559
QY 1177 CTTCAATCAAGATTGCAACCCCGAAACACCTGACTCCAAAGTTCGTATGGTTATCATCA 1236
Db 1560 CTTCAATTAAGATTTGCTCCAGCGGAAGCACCATGCTAAAGTGAGGATGGTGAATATCA 1619
QY 1237 CTGGACCGCCAGAGGCCCAATTTCAAGGCTCAGGGAAGAAATCTATGGCAAACTCAAGGAG 1296
```

```
Db 1620 CTGACACACAGAGGCTCAGTTCAAGGCTCAGGGAAGATTTATGAAAAAATTAAGAAG 1679
QY 1297 AGAACTTCTTTGGTCCCAAGGAGAGTGAAGCTGGAGACCCACATACGTGTGCCAGCAT 1356
Db 1680 AAACTTTGTTAGTCTTAAGAAGAGGTGAACCTTGAGCTCATATCAGAGTGCCTCCT 1739
QY 1357 CAGCAGCTGCGCGGTCAITTTGGCAAGGTGGAAGAAACGGTGAACGATTCAGAAATTTGA 1416
Db 1740 TTGCTGCTGCGCAGAGTATTGGAAAAAGGAGGCAAAACGGTGAATGAATTCAGAAATTTGT 1799
QY 1417 CGCAGCTGAGGTGGTAGTACCAAGAGACAGACCCCTGTATGAGAACGACAGGTCTATCG 1476
Db 1800 CAAGTGCAGAAGTGTGTGTCCTCGTGACCCAGACACCTGATGAGAATGACCAAGTGGTTG 1859
QY 1477 TGAATAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGAGATCC 1536
Db 1860 TCAAAATACTGGTCACTTCTATGCTTGCCAGGTTGCCAGAGAAAAATTCAGAAATTC 1919
QY 1537 TGCCCGAGGTTAAGCAGCAGCATCAGAAAG 1565
Db 1920 TGACTCAGGTAAGCAGCAGCACCACCAACAG 1948

RESULT 12
US-09-480-884A-175
; Sequence 175, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 175
; LENGTH: 4181
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(4181)
; OTHER INFORMATION: n=A,T,C or G
US-09-480-884A-175

Query Match 34.6%; Score 672.6; DB 4; Length 4181;
Best Local Similarity 66.6%; Pred. No. 5.4e-186;
Matches 979; Conservative 0; Mismatches 484; Indels 6; Gaps 1;

QY 103 GGGTTAGAGCGCGGAAAAATTCAGAAATTTCCACCCAGCTCCGATCGGAGTAC 162
Db 480 GGCAAGGATTCGGAACCTTCAGATACGAAATATCCGCTCATTTACAGTGGGAGTGC 539
QY 163 TGGACAGCTGCTGCTCAGTATGTTACAGTAGAAGCTGTGAGCAAGTGAACACCGAGA 222
Db 540 TGGATAGTTTACTAGTCCAGTATGGAGTGGTGGAGAGCTGTGAGCAAGTGAACACTGACT 599
QY 223 GTGAGACGGCAGTGGTGAATGTCACTATTCCAAACCGGAGCAGACCCAGGCAAGCCATCA 282
Db 600 CGGAAACTGCAGTTGTAAATGTAACTTATCCAGTAAAGCAAGCTAGACAGCACTAG 659
QY 283 TGAAGCTGAATGGCCACAGTGTGAGAAACCATGCGCTGAAGTCTCTCTACATCCCGCATG 342
Db 660 ACAAACCTGAATGGATTTTCAGTTAGAGAAATTTCACTTTGAAAGTAGCTTATATCCCTCATG 719
QY 343 AGCAGATAGCACAGGAGACCTGAGAAATGGGCGCCGAGGGGCTTTGGCTCTCGGGGTACG 402
```

Db 1080 TCCCTTGAAGATTAGCTCATAATAACTTTGTGGACGCTCTATTGGTAAGAAGAA 1139
QY 763 GGAACCTGAAGAGGTAGACAGATACCGAGACAAAAATCACCATCTCTCGTTGCAAG 822
Db 1140 GAAATCTTAAANAATTGAGCAAGACACAGACATAAATCACGATATCTCCATGTCAGG 1199
QY 823 ACCTTACCTTTTACAACCTCGAGAGGACCATCTGTGAAGGGGCCCATCGAGAAATGTT 882
Db 1200 AATGACGCTGTATAATCCAGAACGCACTATTACAGTTAAAGGCAATGTTGAGACATGTG 1259
QY 883 GCAGGCGGAGGAGGAATATGAAGAAAGTTCCGGAGGCTATGAGAAATGATGCTG 942
Db 1260 CCAAGCTGAGGAGGAGATCATGAAGAAATCAGGGAGTCTTATGAANAATGATATGCTT 1319
QY 943 CCATGA-----GCTCTACCTGATCCCTGCGCTGAACCTGGCTGCTAGGCTTTTTC 996
Db 1320 CTATGAATCTTCAAGACATTTAATCTCTGATTAATCTGAACGCGCTTGGGTCTGTTC 1379
QY 997 CAGCTTCATCCAGGCGAGTCCCGCGCTCCAGAGCGGTTACTGCGGGCTCTCTCCTATA 1056
Db 1380 CACCCACTTCAGGATGCCACTCCACTCAGGGGCCCTTCAGCCATGACTCTCCCT 1439
QY 1057 GCTCCTTTATCAGGCTCCAGAGGAGATGGTGCAGTCTTTATCCCGCCAGGCGAG 1116
Db 1440 ACCGCGAGTTTGAGCAATCAGAAACGAGACTGTTCATCGTTTATCCAGCTCTATCAG 1499
QY 1117 TGGCGGCATCATCGGCAAGAGGGGAGACATCAACACAGCTCTCCCGGTTTGCAGCG 1176
Db 1500 TCGGTGCCATCATCGGCAAGAGGGCCAGACATCAACAGCTTTCTCGTTTGTGAG 1559
QY 1177 CTTCCATCAAGATTGCACACCCGAAACACCTGTACTCCAAAGTTGTTATGTTATCATCA 1236
Db 1560 CTTCAATTAGATTCTCCAGCGAAGCACCAGATGCTAAAGTGAGGATGTTGATTATCA 1619
QY 1237 CTGACCCGCGAGGCCCAATTCAGGCTCAGGAGAGATCTATGCGAACTCAGAGG 1296
Db 1620 CTGGACCAACAGGCTCAGTTCAAGGCTCAGGGAAGAAATTTATGGAAATAAAGAAG 1679
QY 1297 AGAATCTTTTGTGCTCCAGGAGGAGTGAAGCTGGAGACCAACATAGCTGTGCCAGCAT 1356
Db 1680 AAACTTTGTTAGTCTTAAAGAGAGGTGAACCTTGAAGCTCATATCAGAGTGCATCT 1739
QY 1357 CAGCAGCTGGCGGGTCAATTTGCAAGGTGGAAGAACCGTGAAACGAGTTGCAAGATTGA 1416
Db 1740 TTGCTGCTGGCAGAGTTATTGGAAGAGGAGGCAAAACGGTGAATGAATTCAGAAATTTG 1799
QY 1417 CGGCGAGCTGAGTGTAGTATCAAGAGACGACACCCCTGATGAGAACGACGAGTCAATCG 1476
Db 1800 CAAAGTGCAAGTTGTTGTCCTCGTGACGACACACCTGATGAGAATGACCAAGTGGTTG 1859
QY 1477 TGAATATCATCGACATTTCTATGCGAGTCAGATGGCTCAACGGAAGATCCGAGACATCC 1536
Db 1860 TCAAAATAACTGCTCATCTTATGCTTCCAGGTTGCCCAGAGAAATTCAGGAATTC 1919
QY 1537 TGGCCCGAGTTAAGCAGCAGCATCAGAAG 1565
Db 1920 TGACTCAGGTAAAGCAGCACCACACAG 1948

RESULT 11
US-09-643-597-175
; Sequence 175, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Rongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 175
; LENGTH: 4181
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3347)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3502)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3506)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3520)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3538)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3549)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3646)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3940)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3968)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3974)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4036)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4056)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4062)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4080)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4088)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4115)
; OTHER INFORMATION: n=A,T,C or G
; OTHER INFORMATION: n=A,T,C or G
; US-09-643-597-175

Query Match 34.6%; Score 672.6; DB 4; Length 4181;
Best Local Similarity 66.6%; Pred. No. 5.4e-186;
Matches 979; Conservative 0; Mismatches 484; Indels 6; Gaps 1;
QY 103 GGTTAGGAGCGGAAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTAC 162
Db 480 GGCAAGAGATTCCGAAACTTCAGATACGAAATATCCCGCTCATTTACAGTGGAGGTGC 539

523 CCATCCGCAACATCACAAACAGACCAGTCCAAAGATAGACGTGCATAGGAAGAGAACG 582
|||||
900 CCATTGGAAACATCACAAACAGACCAGTCCAAATCGATGTCCACCGTAAAGAAATG 959
583 CAGGTGACGTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCCCTCGCTT 642
960 CGGGGGCTGCTGAGAGTCGATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCGGGTT 1019
643 GTAAGATGATCTTGAGATATATGATTAAGAGGTAAAGGACACACAAACCGCTGACGAGG 702
1020 GTAAGTCTATTCTGGAGATTATGATAAGGAAGCTCAAGATATAAAATTCACAGAAGA 1079
703 TTCCCTCTGAGATCTCGCCCATATTAATCTTTAGGGCTCTCAATGGCAGGAAGAC 762
1080 TCCCTTTGAGATTTAGCTCATTAATTAATCTTTGAGCGCTTTAATGGTAAAGAGGAA 1139
763 GGAACCTGAAGAGGTAGACAGATACCGAGACAAATAATCACCATCTCTCTGTTGCAAG 822
1140 GAAATCTTAAAAAATTTGACCAAGACACAGACACTAAATCAGATATCTCCATTGAGG 1199
823 ACCTTACCTTTACACCCCTGAGAGGACCATCACTGTGAAGGGGCCCATCGAGAAATGTT 882
1200 AATTGACGCTGTAAATCCAGAACGCATTAACAGTTAAAGGCAATGTTGAGACATGTG 1259
883 GCAGGGCCGACGAGGAATAATGAAGAAAGTTGCGGAGGCTATGAGAATGATGTGCTG 942
1260 CCAAGCTGAGGAGGAGATCATGAAGAAATTCAGGGAGTCTTATGAAAAATGATTTGCTT 1319
943 CCATGA-----GCTCTCACTGATCCCTGGCCCTGAACTGGCTGCTGTAGGCTTTTCC 996
1320 CTATGAATCTTCAAGCACATTAATCTCTGATTAATCTGAACGCTTGGCTGTGTTCC 1379
997 CAGCTTCATCAGGCGAGTCCCGCGCTCCAGAGCGGTTACTGGGGCTGCTCCCTATA 1056
1380 CACCCACTTCAGGGATGCCACCTCCACCTCAGGGCCCCCTTCAGCCATGACTCCTCCCT 1439
1057 GCTCTCTTTATGAGGCTCCCGAGCAGGAGATGTGCAAGTGTGAGGTGTTATCCCGCCGAGCAG 1116
1440 ACCCGAGTTTGAGCAATCAGAAACGGAGACTGTTATCATGATTTATCCCGACTCTATCAG 1499
1117 TGGGCGCCATCATCGGCAAGAGGGGAGCACATCAAAACAGCTCTCCCGGTTTGGCAGCG 1176
1500 TCGGTGCCATCATCGGCAAGCGGCGAGCACATCAAGCAGCTTTCTCGTTTGTCTGGAG 1559
1177 CTTCCATCAAGATTTGACCAACCGAACACCTCACTCAAGTTCGTAAGTTGATGTTATCATCA 1236
1560 CTTCAATTAAGATTGCTCCAGCGGAAGACACAGATGCTAAAGTGAGGATGGTGAATATCA 1619
1237 CTGGACCGCAGAGGCCCAATTCAGGCTCAGGGAAGATCTATGCAAACTCAAGGAGG 1296
1620 CTGGACCCACAGAGGCTCAGTTCAAGGCTCAGGGAAGATTTATGGAATAATTAAGAAG 1679
1297 AGAATCTTTGTTGTTCCAGAGGAAGTGAAGCTGGAGACCCACATACGTTGTCAGCAT 1356
1680 AAAACTTTGTTAGTCTTAAAGAGAGGTGAAACTTGAAGCTCATATCAGATGSCCATCCT 1739
1357 CAGCAGCTCGCGGGTCATTTGCAAGAGTGGAAACCGGTGAACGATGACGAGTTGACAGATTTGA 1416
1740 TTGCTCTGGCAGAGTTATGGAAGAGGAGGCAAAACCGGTGAATGAATTCAGAAATTTGT 1799
1417 CGGCACTGAGGTGTTAGTACCAAGAGACAGACCCCTGATGAGACGACCGAGTCAATCG 1476
1800 CAAGTSCAGAGTTGTTGTCCTCTGTCGACGACACCTGATGAGATGACCAAGTGGTGTG 1859
1477 TGAATAATCATCGACATTTCTATGTCAGTCAAGTGGCTCAACGGAAAGATCCGAGACATCC 1536
1860 TCAAAATAAATGCTGCTCTATGCTTGCAGGTTGCCAGGTTGCCAGAGAAAAATTCAGGAAATTC 1919
1537 TGGCCAGAGTTAAGCAGGACATCAGAG 1565
1920 TGAATCAGGTAAGCAGGACCAACACAG 1948

RESULT 10

US-09-899-651-4
; Sequence 4, Application US/09899651
; Patent No. 6576756

GENERAL INFORMATION:

; APPLICANT: Chen, Yao-Tsung
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; TITLE OF INVENTION: Associated
; FILE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/899,651
; PRIOR FILING DATE: 2001-07-06
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 4
; LENGTH: 4159
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-899-651-4

Query Match 34.6%; Score 672.6; DB 4; Length 4159;
Best Local Similarity 66.8%; Pred. No. 5.4e-186;
Matches 979; Conservative 0; Mismatches 484; Indels 6; Gaps 1;

QY 103 GGTTAGGAGCCGGAAAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGAAAGTAC 162
Db |||||
QY 480 GGCAGAGATTGCGAACTTCAGATACGAATATCCCGCTCATTTACAGTGGAGGTGC 539
Db |||||
QY 163 TGAACAGCTGCTGGCTCAGTATGGTACAGTACAGAACTGTGAGCAAGTGAACACCGAGA 222
Db |||||
QY 540 TGGATAGTTTACTAGTCCAGTATGGAGTGGTGGAGAGCTGTGAGCAAGTGAACACTGACT 599
QY 223 GTGAGACGCGAGTGTGAATGTCACTATTTCCAACTGAGAGAGACGAGCAAGCATCA 282
Db |||||
QY 600 CGAAACTCAGTGTAAATGTAACTATTCAGTAAGCAACCAAGCTAGACAGCACTAG 659
QY 283 TGAAGCTGAATGGCCACCACTGGAGAACCATGCTCCCTGAAGGTCTCTACATCCCGATG 342
Db |||||
QY 660 ACAACTGAATGATTTCAAGTTAGAGATTTTCACTTTGAAGTAGCTATATCCCTGATG 719
QY 343 AGCAGATACACAGGACCTGAGATGGGCGCGAGGGGCTTTGGCTCTCGGGGTGACG 402
Db |||||
QY 720 AAATGGCCCGCCAGCAAAAACCCCTTTGACAGCAGCCCGAGGTGCGCCGGGGCTTGGGCGA 779
QY 403 CCGCCAGGGCTCACCTGTGGCAGCGGGGGCCCCAGCAAGCAGCAGCAAGTGAATCC 462
Db |||||
QY 780 GGGGCTCTCAGGCGAGGGGTCTCCAGATCCGTATCCAGCAGAAACCATGTGATTGC 839
QY 463 CCCTTCGGCTCCTGTGCCCCACCCAGTATGTGGGTGCCATTATTTGCCAAGGAGGGGCCA 522
Db |||||
QY 840 CTCTGCGCTGCTGTTCCACCCCAATTTGTTGGAGCCATCATAGGAAAGAGGTGCCA 899
QY 523 CCATCGCNAATCACAACAGACCCAGTCCAGATAGACGTGATAGGAGGAGAACG 582
Db |||||
QY 900 CCAATTCGAACATCACCAACAGACCCAGTCTTAAATCGATGTCCACCGCTTAAAGAAATG 959
QY 583 CAGGTGACGTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCTCGGCTT 642
Db |||||
QY 960 CGGGGCTGCTGAGAGTGCATTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCGGGCTT 1019
QY 643 GTAAGATGATCTTGAGATTTATGCAATAAGAGGCTTAAGGACACCAACCGCTGACGAGG 702
Db |||||
QY 1020 GTAAGTCTATTCTGAGATTTATGCAATAGGAAGCTCAAGATATAAAATTCAGAGAGAGA 1079
QY 703 TTCCCTCAAGATCTCGGCCAATAATACTTTGTAGGGCTCTCATTCAGCAAGGAAGAC 762

283	QY	TGAAGCTGAATGCCACCACTGTTGGAGAACCATGSCCTGAAGGCTCTCTACATATCCCCGATG	342
410	Db	ACAAACTGAATCGATTTCAGTTTAGAGAAATTCACCTTGAAGTAGCCTATATCCCTGATG	469
343	QY	AGCAGATAGCAACGGGACCTGAGAAATGGCGCGCGAGGGGCTTTGGCTCTCGGGGTACGC	402
470	Db	AAACGGCGGCCACAGCAAAAACCCCTTGCACAGCCCCGAGGTGCGCGGGGGCTTTGGGCAGA	529
403	QY	CCCGCCAGGGCTCACCTGTGGGACGCGGGCGCCAGCCAAAGCAGCAGCAAGTGGACATCC	462
530	Db	GGGGCTCTCAAGGCAGGGGTCTCCAGATTCGGTATCCCAAGCAGAAACCAATGTGATTGC	589
463	QY	CCCTTCGGCTCTCGTGCCTCCACCCAGCATATGTGGTGCCATTATTTGGCAAGAGGGGGCA	522
590	Db	CTCTGCGCTCTGTTTCCCACCCAAATTTGTTGGAGCCATCATAGMAAAGAAGTGCACA	649
523	QY	CCATTCGCGAACTACAAAAACAGACCCAGTCCAGATAGACGTGCATAGGAGAGAGAACG	582
650	Db	CCATTCGGAACATCAACAAACACAGCCAGTCTAAATCGATGTCCACCGTAAAGAAAATG	709
583	QY	CAGTGCAGCTGAAAAGGCCATCAGTGTGCATCCACCCCTCAGGGCTCTCTCCCTCGGTT	642
710	Db	CGGGGGTCTGTGAGAGTCAATCTCTCTACTCTGTGAGGCACCTCTGCGGCTT	769
643	QY	GTAAGATGATCTTTGGAGATTATGCATAAAGAGGCTTAAGACACCAAAACGGCTGACGAGG	702
770	Db	GTAAGTCTATTCTGGAGATTATGCATAAAGGAAGCTCAAGATATAAAATTCACAGAAGAGA	829
703	QY	TTCCCTTGAAGATCTCGGCCCTAATAAATCTTTGTAGGGCTCTCATTTGGCAAGGAGGAC	762
830	Db	TCCCTTTGAAGATTTTAGCTCATATAACTTTGTTGGAGCTCTTATTTGTAAGAAGAGAA	889
763	QY	GGAACTTGAAGAGTTAGACAAGATACCGAGACAAAATCACCATCTCTCTGTTGCAAG	822
890	Db	GAAATCTTAAAAAATTTGACAGAGACACAGACACTAAATCAGATATCTCCATTGCAGG	949
823	QY	ACCTTACCTTTTACAACCTCTGAGAGGACCATCTCTGTGAAGGGGGCCATCGAGAAATGTT	882
950	Db	AAATTGACGCTCTATAATCCAGAGCGCACTATTACAGTTAAAGGCAATGTTGACACATGTG	1009
883	QY	GCAGGGCGGACAGGAAATTAATGAAGAAATTTGGGAGGCTATGAGATGATGTGGCTG	942
1010	Db	CCAAAGCTGAGGAGGAGATCATGAAGAAAATACGGGAGTCTTATGAATAATGATTTGCTT	1069
943	QY	CCATGA-----GCTCTCACTGATCCTCGCTGACCTGCTGTGTAGTCTTTTCC	996
1070	Db	CTATGAATCTTCAAGACATTTAATCTCGATTAATCTGAATTAATCTGAACGCTTTGGGTCTG	1129
997	QY	CAGCTTCAATCCAGCGCAGTCCCGCGCTCCACAGCAGCGTTACTCGGGGCTGCTCCCTATA	1056
1130	Db	CACCACTTCAGGGATGCCACTCCACCTCAGGGCCCCCTTCAGCCATGACTCTCTCCCT	1189
1057	QY	GCTCTCTTATGAGGCTCCGAGCAGAGATGTGTGAGGCTTATGAGATGATGTGGCTG	1116
1190	Db	ACCCGCACTTTGACCAATCAGAAACGGAGACTGTTCATCTGTTTATCCAGCTCTATCAG	1249
1117	QY	TGGGCGCATCATCGGCAAGAGGGCAGACATCAAAACAGCTCTCCCGGTTTGGCCAGCG	1176
1250	Db	TCGGTGCCATCATCGGCAAGCAGGGCCAGACATCAAGCAGCTTTCTCGCTTTGCTGAG	1309
1177	QY	CCTCCATCAGATTTGCACCCCGAAACACCTGACTCCAAAGTTTGTATGGTTATCATCA	1236
1310	Db	CTTCAATTAAGATTGCTCCAGCGGAAGCACCCAGATGCTAAAGTGAGGATGGTGAATATCA	1369
1237	QY	CTGGAACGCGCAGAGCCCAATTCAGGCTCAGGAGAGAACTATGCGCAAACTCAGAGAGG	1296
1370	Db	CTGGAACCAACAGAGCTCAGTTCAAGGCTCAGGAGAGAAATTTATGGAATAATTAAGAAG	1429
1297	QY	AGAACTCTTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCGCAGCAT	1356
1430	Db	AAAATCTTGTAGTCTTAAAGAGAGGTGAACCTTGAAGCTCATATCAGAGTGCCATCT	1489
1357	QY	CAGCAGCTGGCCGGGTCAATGSCAAAGGTGGAAAACCGGTGAAACGAGTTTCAGAAATTGA	1416

```

Db      1490  TTGCTGCTGCAGATTATTTGAAAAGGAGGCAAAACGGTGAATGAATCACTTCAGAAATTGT 1549
               |||
QY      1417  CGGCAGCTGAGGTGGTAGTACCAAGACACGACCCCTGATGAGAACGACACGCTCATCG 1476
               |||
Db      1550  CAAGTCAGAAAGTTGTTGCTCCCTCGTGACCAGACACCTGATGAGAATGACCAAGTGGTGG 1609
               |||
QY      1477  TGAATCATTCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCC 1536
               |||
Db      1610  TCAAAATAACTGGTCACTTCTATGCTTGCAGGTTGCCAGAGAAAATTCAGGAAATTC 1669
               |||
QY      1537  TGGCCCAAGTTAAGCAGCAGCATCAGAAG 1565
               |||
Db      1670  TGACTCAGTTAAGCAGCACCACCAACAG 1698
               |||

RESULT 9
US-09-061-709-4
; Sequence 4, Application US/09061709B
; Patent No. 6297364
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tsung
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuch, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
; FILE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/061,709B
; CURRENT FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 4
; LENGTH: 4159
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-061-709-4

Query Match      34.6%;      Score 672.6; DB 3;      Length 4159;
Best Local Similarity 66.6%;      Pred. No. 5.4e-186;
Matches 979; Conservative 0; Mismatches 484; Indels 6; Gaps 1;

QY      103  GGGTTAGAGCGCGGAAAATTCAAATCCGAATATTTCCACCCAGCTCCGATCGGAGTAC 162
Db      480  GCAAAGGATTTCGGAACCTTCAGATACGAAATATCCCGCTCATTTACAGTGGGAGGTGC 539
QY      163  TGGACAGCCTGCTGCTCAGTATGTTACAGTAGAGAACTGTGAGCAAGTGAACACCGAGA 222
Db      540  TGGATAGTTTACTGTCAGTAAGAGTGGTGGAGAGCTGTGAGCAAGTGAACACTGACT 599
QY      223  GTGACACGCGCAGTGGTGAATGTCACTATTTCCAAACGGGAGCAGACCGAGAGCCATCA 282
Db      600  CGGAACCTGCAGTTGTAATGTAACTATTCAGTAAGCAACCAAGCTAGACAAGCACTAG 659
QY      283  TGAAGCTGAATGGCCACGAGTTGGAGAACCATGCCCCTGAAGTCTCTCTACATCCCGCATG 342
Db      660  ACAAACTGAATGGATTTTCAGTTAGAGAAATTTCCACCTTGAAGTAGGCTATATCCCTGAATG 719
QY      343  AGCAGATACACAGGAGACCTAGAAATGGCGCGCAGGGGGGCTTTGGCTCTCGGGGTGAGC 402
Db      720  AATGGCCGCCAGCAAAACCCCTTGCAGCAGCCCGGAGGTGCGCGGGGGCTTTGGCAGA 779
QY      403  CCGCGCAGGGCTCACTGTGTGGCAGCGGGGGCCCCAGCCAGCAGCAGCAAGTGGGATCC 462
Db      780  GGGGCTCTCAAGGCAGGGGTCTCCAGGATCCGATATCCAAAGCAGAAACCATGTGATTTGC 839
QY      463  CCCTTCGGCTCTCTGGTCCCAACCCAGTAGTGGGTGSCATATTGTGCAAGAGGGGGCCA 522
Db      840  CTCCTGGCTCTGGTGTCCCAACCCAAATTTGTTGGAGCCATCATAGAAAAGAGGTGCCA 899

```



```

; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 347
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-643-597-347

Query Match 34.6%; Score 674.2; DB 4; Length 1740;
Best Local Similarity 66.7%; Pred. No. 1.1e-186;
Matches 980; Conservative 0; Mismatches 483; Indels 6; Gaps 1;

QY 103 GGGTTAGGACCGGAAATTCAAATCGAAATATTCACCCCGAGCTCCGATGGGAAGTAC 162
DB 230 GGCAAGAGATTCGAAACTTCAGATACGAATATCCCGCTCATTTACAGTGGGAGGTGC 289
QY 163 TGACAGCCCTGCTGGCTCAGTATGGTACAGTACAGTCTGTAGCAAGTGAACACCGAGA 222
DB 290 TGGATATGTTTACTAGTCCAGTATGGAGTGTGGAGAGCTGTAGCAAGTGAACACCTGACT 349
QY 223 GTGAGACGGCAGTGGTGAATGTCACTATTCACACCGGGAGCAGACCGAGCAAGCCATCA 282
DB 350 CGGAACCTGAGTGTGAATGTAAATGTAACCTATTCAGTAGGACCAAGCTAGACAGCACTAG 409
QY 283 TGAAGCTGAATGGCCACAGTTGGAGAACCATGCGCTTGAAGGTCCTCATCATCCCGATG 342
DB 410 ACAAACTGAATGGATTTTCAGTTAGAGAAATTCACCTTTGAAAGTAGCCTATATCCCTGATG 469
QY 343 AGCAGATAGACAGGACCTGAGAAATGGCGCCGAGGGGCTTTGGCTCTCGGGTCCAGC 402
DB 470 AAACGGCGCCCGACAAACCCCTTGCAGCAGCCCGAGGTCGCGGGGGCTTGGGCGAGA 529
QY 403 CCGCCAGGGCTCACCTGTGCGCGGGGGCCCGAGCAAGCAGCAGCAAGTGGACATCC 462
DB 530 GGGGCTCTCAAGCAGGGGTCTCCAGATCCGATATCCAGCAGAAACCATGTGATTTGC 589
QY 463 CCCTTCGGCTCCTGGTCCCAACCCAGATATGFGGTGCGATATTTGGCAAGGAGGGGCCA 522
DB 590 CTCTGGCGCTGCTGGTTCGCCCAATTTGTTGGAGCCATCATAGGAAAGAAAGTGCCCA 649
QY 523 CCATCCGCAACATCAGAAACAGACCCAGTCCAGATAGAGCTCATAGGAAGGAGAACG 582
DB 650 CCATTCGGAACATCAGCAACAGACCCAGTCTAAATCGATGTCACCCGTAAAGAAATG 709
QY 583 CAGGTGAGCTGAAAGAGCCATCAGTGTGCACTCCACCCCTGAGGGGTGCTCTCCGCTT 642
DB 710 CGGGGCTGCTGAGAGTGAATGATCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 769
QY 643 GTAAGATGATCTTGAGATTTATGCAATAAGAGGTAAAGACACCAAAACGGCTCAGCAGG 702
DB 770 GTAAGTCTATTTCTGGAGATTTATGCAATAAGAGGTCTCAAGATATAAAATTCACAGAAGA 829
QY 703 TTCCCTGAAGATCTGCGCCCAATATATCTTTGAGGGCTCTCATTTGGCAAGGAGGAC 762
DB 830 TCCCTTGAAGATTTAGCTCATATAATCTTTGAGACGCTTATTTGGTAAAGAGGAA 889
QY 763 GGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAAATCACCATCTCTCTCTCTCTCTCT 822
DB 890 GAAATCTTAAATAATTGAGCAAGACACAGACACTAAATCACGATATCTCCATTTGACG 949
QY 823 ACCTTACCTCTTTACAACTCTGAGAGGACCATCATCTGTGAAGGGGGCCCATCGAGATTTGT 882
DB 950 AATTGACGCTGTATAATCCAGAACCCACTATTTACAGTTTAAAGGCAATTTGAGACATGTG 1009
QY 883 GAGGGCCGAGCAGAAATATGAGAGAGTTGCGGGGCGCTTATCAGATATGATGGCTG 942

```

Best Local Similarity 83.5%; Pred. No. 0;
Matches 1556; Conservative 0; Mismatches 275; Indels 33; Gaps 6;

```
QY 108 AGGAGCGGAAATTCGAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTACTGGAC 167
    |||||
Db 365 AGGAGTCGGAATAATACAGATCCGCAATATTCACCTCAGCTCCGATGGGAAGTCTAGAT 424
    |||||
QY 168 AGCTGTCTGGCTCAGTATGTPACAGTAGAGAACTGTGAGCAAGTGAACACCGAGAGTGAG 227
    |||||
Db 425 AGCTGTCTGGCTCAGTACGGTACAGTGGAACTGTGAGCAAGTGAACACTGAAAGTGAG 484
    |||||
QY 228 ACGCAGTGTGTAATGTCACTATTCACACCGGAGCAGACCGGCAAGCCATCATGAAG 287
    |||||
Db 485 ACAGCGTGTGTCACGTCACCTACTCTTAAACCGGAGCAGACCGGCAAGCTATCATGAG 544
    |||||
QY 288 CTGAAATGGCACACAGTTGGAGAACCATGCTGTAAGGTCTCTTACATCCCGATGAGCAG 347
    |||||
Db 545 CTAAATGGCCATCACTGGAGAACCATGCTGTAAGGTCTCTTACATACCTGATGAGCAG 604
    |||||
QY 348 ATAGCACAGGACCTGAGAAATGGCGCGCGAGGGGGCTTTGGCTCTCGGGGTCAAGCCGCG 407
    |||||
Db 605 ATACACAAAGTCTGAGAAATGGCGCTGCTGGAGGCTTTGGGTCTCGGGGCGAGCCCGG 664
    |||||
QY 408 CAGGGTCTCACTGTGGCAGCGGGGGCCCGACGCAAGCAGCAGCAAGTGGACATCCCCCTT 467
    |||||
Db 665 CAAAGGTGCGCCGTGGCAGCAGGGGCTCCAGCCAAAGCAGCAGTGGACATCCCTCTC 724
    |||||
QY 468 CGGCTCTGTGTGCCACCCAGTATGTGGTGCCATTTATGGCAAGAGGGGGCCACCATC 527
    |||||
Db 725 CGGCTCTGTGTGCTACGCGAGTATGTAGGGGCTATCATTTGGCAAGAGGGGTGCGCCATC 784
    |||||
QY 528 CGCAACATCAAAACAGACCCAGTCCAGATAGACGTGCATAGGAGGAGAGCGCAGGT 587
    |||||
Db 785 CGAAACATCAAAACAGACCGAGTCCAAATAGACGTGCATAGGAGGAGAAATGGGGC 844
    |||||
QY 588 GCAAGTGAATAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCCTCCGCTTGTAAAG 647
    |||||
Db 845 GCTCGGAGAAAGCCATCAGCTGTCATTCAACCCCTGAAGGCTGCTCCTCGGCTGCAAG 904
    |||||
QY 648 ATGATCTTGAGATATGCAATAAGAGGCTTAAGACACCAAAAGCGGTGACGAGTTTCC 707
    |||||
Db 905 ATGATCTTGAGATATGCAACAGGAGGCAAGGACACCAAAACGGCGAGTGAAGTTTCC 964
    |||||
QY 708 CTGAAGATCTGGCCCATATAACTTTGTAGGGGCTCTCATTTGCAAGGAGGACGCAAC 767
    |||||
Db 965 CTGAAGATCTGGCTCATACAACTTCTGCGGGGACTCATTTGGCAAGGAGGCGGCAAC 1024
    |||||
QY 768 CTGAAGAGGTGAGCAAGATACGAGACAAATAATCACCATCTCCTCGTTGCAAGACCTT 827
    |||||
Db 1025 CTGAAGAGGTGAGCAGGACACAGAGACGAAGATCACCATCTCATCGCTCCAGGACCTC 1084
    |||||
QY 828 ACCCTTTACAACTTGAAGGACCATCAGTGAAGGGGCCATCGAGAAATGTTGAGG 887
    |||||
Db 1085 ACCTCTTAACCTGAGAGGACCATCAGTGAAGGGGCCCATTTGAGAACTGTTGAGG 1144
    |||||
QY 888 GCGAGCAGGAATAATGAAGAAGTTTCGGAGGCGCTATGAGAAATGATGTGGCTGCCATG 947
    |||||
Db 1145 GCGAGCAGGAGATCATGAGAAAGTTTCGAGAGGCTTACGAGAACGAGTGGCGGCAATG 1204
    |||||
QY 948 AGC-----TCTCACTGATTCCTCGGCTGAACTCTGGCTGCTAGGTCTTTTCCAGCT 1001
    |||||
Db 1205 AGCTTGAGTCCCACTCATCTCCCTGGCTTAACTGGCTGCTAGGTCTCTTTCCAGCT 1264
    |||||
QY 1002 TCATCCAGCGAGTCCCGGCTCCCGAGGAGTACTGGGCTGCTCCCTTAGCTCC 1061
    |||||
Db 1265 TCATCCAGCGCTGCTCCCTCTCTCCAGAGGTGCTCACTGGGCTGCTCCCTTAGCTCC 1324
    |||||
QY 1062 TTTATGAGGCTCCCGAGCAGGAGATGTTGACAGGTGTTATCCCGCCAGGCGAGGGC 1121
    |||||
Db 1325 TTATGAGGCTCCCGAGCAGGAGATGTTACAGTGTTCATCCCGCCAGGCTGTGGGC 1384
    |||||
QY 1122 GCATCATCGGCAGAGGGGAGCACAATCAACAGCTCTCCCGGTTTGCCAGCGCTCC 1181
    |||||
```

```
Db 1385 GCCATCATTTGGCAAGAGGGCCAGCACATCAAAACAATCTCTCCGTTTTCGCCAGCGCTCC 1444
QY 1182 ATCAAGATTGCACACCCGAAACACCTGACTCCAAAGTTCGTATGTTATCATCACTGGA 1241
    |||||
Db 1445 ATCAAGATTGCTCCACCAGAAACACCTGACTCCAAAGTTCGAATGGTCTCATCACTGGA 1504
    |||||
QY 1242 CCGCAGAGGCCCAATTCAGGCTCAGGGAAGAAATCTATGGCAAACTCAAGGAGGAGAAC 1301
    |||||
Db 1505 CCCCCAGAGGCTCAGTTCAAGGCTCAGGGAAGAATTTATGGCAAACTAAAAAGAGAAT 1564
    |||||
QY 1302 TTCTTTGGTCCCAAGAGGAGTGAAGCTGAGACCCACATACGTGTGCCAGCATCAGCA 1361
    |||||
Db 1565 TTCTTTGGTCCCAAGAGGAGTGAAGCTGAGACCCACATACGGGTTCCGGCTTCAGCA 1624
    |||||
QY 1362 GCTGGCGGGTCTATTTGGCAAAAGGTGAAACCGGTGAACCGAGTTGCAGAAATTTGACGGCA 1421
    |||||
Db 1625 GCCGGCGGCTCATTCGGCAAAAGCGGCAAAACGGTGAATGAGCTGCAGAACTTGACTGCA 1684
    |||||
QY 1422 GCTGAGTGTGTAGTACCAAGAGCAGACCCCTGTAGTAGAAACGACCAAGTCTATCGTAAA 1481
    |||||
Db 1685 GCTGAGTGTGTAGTGCACAGAGACCCAGACCCCGGTGAGAACGACCAAGTCTATTTAAG 1744
    |||||
QY 1482 ATCATCGGACATTTCTATGCCAGTCAAGTGGCTCAACGGGAAGATCCGAGACATCTCGGCC 1541
    |||||
Db 1745 ATCATCGGACATTTCTATGCCAGTGGCTCAGCGGAAGATCCGAGACATCTCTGGCT 1804
    |||||
QY 1542 CAGTTAAGCAGCAGCATCAGAAAGGACAGAGTAACAGGCCCGCAGGCAAGGAGGAAAGTGA 1601
    |||||
Db 1805 CAAAGTTAAGCAACAGCAGCAGCAAGAGGACAGAGCAACTGSCCCAGGACCGGAGGAGTGA 1864
    |||||
QY 1602 -CAGGCCCTCTCTGTCCTTNGAGTCCAGGACAAACAAAGGGCAGAA----- 1647
    |||||
Db 1865 CCCCCCGGCTCTCTGTCCTTCCAGATTTGGCTCAGAGAGGAGAACACAGAACTGGAGGG 1924
    |||||
QY 1648 -----ATCGAGAGTGTCTCTCCCGCAGGCTGAGAAATGAGTGGGAATCCGGAC 1699
    |||||
Db 1925 GCGGTTGAGGGCGCGGTGTGTTTTCCAGAGGCTGAGATGAGTGGGAATCAGGG-C 1983
    |||||
QY 1700 ACNWTGGCGGGGTGTAGATCAGGTTTGGCCCACTTCAATTCAGAAAGATGTTCCAGTGAGG 1759
    |||||
Db 1984 ATTGCGGCTGCTGGAGATCAGTTTGCACACTGTATTGAGAAACAATGTTCCAGTGAGG 2043
    |||||
QY 1760 AACCTGTATCTACGCCCCCAACCCACCCAACTTGGCCCAACACTGTTGCCCCCTCG 1819
    |||||
Db 2044 ATCTGTATCTCTCGCCCCCAA--TTGAGCAGCTGGCCACAGCCCACTTTGSAATAT 2101
    |||||
QY 1820 GGTGTGAGAAATNTAGCGCAAGGCACTTTTAAACGTGGATTTTAAAGAGCTCTCCA 1879
    |||||
Db 2102 CACCATTTGCAATCATAGCTTTGGTTGCTTTTAAACGTGGATTTGCT--TGAAGTTCTCCA 2159
    |||||
QY 1880 GGGCCCAACAGAGGGTGGATCACCTCAGTGGGAAGAAAAATAAAATTTCTTCAGGT 1939
    |||||
Db 2160 GCCTCCATGGAAGGATGGGTGATCCAGTGGGGAAGAAATAAAATTTCTTCAGGT 2219
    |||||
QY 1940 TTTA 1943
    |||||
Db 2220 TTTA 2223
    |||||
```

RESULT 6

US-09-643-597-347
; Sequence 347, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.

; CURRENT APPLICATION NUMBER: US/09/061.709B
; CURRENT FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 5
; LENGTH: 1708
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-061-709-5

Query Match 85.8%; Score 1669.2; DB 3; Length 1708;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1676; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 268 CCAGGCAAGCATCATCAAGCTGAATGCGCCACCAAGTTGGAGAACCATGACCTGAAGTCT 327
DB 30 CCGGGGAGCCATCATGAAGCTGAATGGCCACCAAGTTGGAGAACCATGACCTGAAGTCT 89
QY 328 CTTACATCTCCCGATGAGCAGATAGCACAGGAGCCTGAGAAATGGCGCCGAGGGGGCTTTG 387
DB 90 CTTACATCTCCCGATGAGCAGATAGCACAGGAGCCTGAGAAATGGCGCCGAGGGGGCTTTG 149
QY 388 GCTCTCGGGGTGACCCCGCCAGGCTCACCTGTGGCAGCGGGGGCCCAAGCAGC 447
DB 150 GCTCTCGGGGTGACCCCGCCAGGCTCACCTGTGGCAGCGGGGGCCCAAGCAGC 209
QY 448 AGCAAGTGGACATCCCGCTCTCGCTCTGCTGTCGCCACCAAGATGAGGAGTATGTTG 507
DB 210 AGCAAGTGGACATCCCGCTCTGCTGTCGCCACCAAGATGAGGAGTATGTTG 269
QY 508 GCAAGGAGGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCCAAGATAGACGTGC 567
DB 270 GCAAGGAGGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCCAAGATAGACGTGC 329
QY 568 ATAGGAGGAGACCGAGGTGACGTGAAAGCCATCAAGTGTGACCTCCACCCCTGAGG 627
DB 330 ATAGGAGGAGACCGAGGTGACGTGAAAGCCATCAAGTGTGACCTCCACCCCTGAGG 389
QY 628 GCTGCTCTCTCCGCTTGAAGATGATCTTGAGATTAATGATGAAGAGGCTTAGGACACCA 687
DB 390 GCTGCTCTCTCCGCTTGAAGATGATCTTGAGATTAATGATGAAGAGGCTTAGGACACCA 449
QY 688 AAACGGCTGACGAGGTTCCTTGAAGATCTTGCCCAATTAATCTTTAGGGCGTCTCA 747
DB 450 AAACGGCTGACGAGGTTCCTTGAAGATCTTGCCCAATTAATCTTTAGGGCGTCTCA 509
QY 748 TTGGCAAGCAAGACCGACCTGACAGAGTAGACGACATCCGAGACAAATCAACA 807
DB 510 TTGGCAAGCAAGACCGACCTGACAGAGTAGACGACATCCGAGACAAATCAACA 569
QY 808 TCTCTCTGTTGAAGACCTTACCTTTACACCTGAGAGGACCATCACTGTGAAGGGGG 867
DB 570 TCTCTCTGTTGAAGACCTTACCTTTACACCTGAGAGGACCATCACTGTGAAGGGGG 629
QY 868 CCATCGAATTTGTCAGGGCGGAGCAGGAATTAATGAAGAGTTCGGGAGGCGCTATG 927
DB 630 CCATCGAATTTGTCAGGGCGGAGCAGGAATTAATGAAGAGTTCGGGAGGCGCTATG 689
QY 928 AGAATGATGTGCTGACATGCTCTCACTGATCCCTGCGCTGACCTGGCTGCTGTAG 987
DB 690 AGAATGATGTGCTGACATGCTCTCACTGATCCCTGCGCTGAACTGGCTGCTGTAG 749
QY 988 GTCTTTTCCAGCTTTCATCCAGGCGAGTCCCGCGCTCCAGCAGCGTTACTGGGGCTG 1047
DB 750 GTCTTTTCCAGCTTTCATCCAGGCGAGTCCCGCGCTCCAGCAGCGTTACTGGGGCTG 809
QY 1048 CTCCTATAGTCTCTTTATGAGGTCCCGAGCAGAGATGCTGAGTGTATTATCCCGG 1107
DB 810 CTCCTATAGTCTCTTTATGAGGTCCCGAGCAGAGATGCTGAGTGTATTATCCCGG 869
QY 1108 CCCAGGAGTGGGGCCCATCATCGCAAGAGGGGCGACATCAAAAGCTCTCCCGT 1167
DB 870 CCCAGGAGTGGGGCCCATCATCGCAAGAGGGGCGACATCAAAAGCTCTCCCGT 929

QY 1168 TTGCAGGCGCTCATCAAGATTGCAACCCGAAACACCTGACTCCAAAGTTGCTATGG 1227
DB 930 TTGCAGGCGCTCATCAAGATTGCAACCCGAAACACCTGACTCCAAAGTTGCTATGG 989
QY 1228 TTATCATCACTGGACCGCAGAGGCGCAATTCAGGCTCAGGGAAGAAATCTATGGCAAAAC 1287
DB 990 TTATCATCACTGGACCGCAGAGGCGCAATTCAGGCTCAGGGAAGAAATCTATGGCAAAAC 1049
QY 1288 TCAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTG 1347
DB 1050 TCAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTG 1109
QY 1348 TGCAGCATCAGCAGCTGGCCGGGTCAATGGCAAAAGTGGAAACCGGTGAACAGTTCG 1407
DB 1110 TGCAGCATCAGCAGCTGGCCGGGTCAATGGCAAAAGTGGAAACCGGTGAACAGTTCG 1169
QY 1408 AGAATTTGACGCGCAGCTGAGTGTAGTACCAAGAGACCCAGACCCCTGATGAGAACGACC 1467
DB 1170 AGAATTTGACGCGCAGCTGAGTGTAGTACCAAGAGACCCAGACCCCTGATGAGAACGACC 1229
QY 1468 AGGTCACTCTGAAAAATCATCGGACATTTCTATGCCAGTCAAGTGGTCAACGGAAGATCC 1527
DB 1230 AGGTCACTCTGAAAAATCATCGGACATTTCTATGCCAGTCAAGTGGTCAACGGAAGATCC 1289
QY 1528 GAGACATCTGCCCCAGGTTAAGCAGCAGCATCAGAAGGGACAGAGTAACCCAGGCCCAGG 1587
DB 1290 GAGACATCTGCCCCAGGTTAAGCAGCAGCATCAGAAGGGACAGAGTAACCCAGGCCCAGG 1349
QY 1588 CACGAGGAGAGTACACAGCCCTCCCTGTCCTTNGAGTCCAGGACAAACACGCGGAGAA 1647
DB 1350 CACGAGGAGAGTACACAGCCCTCCCTGTCCTTNGAGTCCAGGACAAACACGCGGAGAA 1409
QY 1648 ATCGAGGTGTCTCTCTCCCGCAGGCTGAGAATGAGTGGGAATCCGGGACACNTGGGC 1707
DB 1410 ATCGAGGTGTCTCTCTCCCGCAGGCTGAGAATGAGTGGGAATCCGGGACACNTGGGC 1469
QY 1708 CCGGCTGTAGATCAGGTTTGCCCACTTGATGAGAAAGATGTTCCAGTCAAGGAACCTCGA 1767
DB 1470 CCGGCTGTAGATCAGGTTTGCCCACTTGATGAGAAAGATGTTCCAGTCAAGGAACCTCGA 1529
QY 1768 TCTNTCAGCCCCCAACACCCACCTTAAACGTTGGATTTTAAAGAGCTCTCCAGGCCCCAC 1887
DB 1530 TCTNTCAGCCCCCAACACCCACCTTAAACGTTGGATTTTAAAGAGCTCTCCAGGCCCCAC 1649
QY 1828 AAATTTAGCGCAAGGCACTTTTAAACGTTGGATTTTAAAGAGCTCTCCAGGCCCCAC 1887
DB 1590 AAATTTAGCGCAAGGCACTTTTAAACGTTGGATTTTAAAGAGCTCTCCAGGCCCCAC 1649
QY 1888 CAAGAGGTGATCACACCTCAGTGGGAGAGAAAAATAAAATTTCTTCAGGTTTTAAAA 1946
DB 1650 CAAGAGGTGATCACACCTCAGTGGGAGAGAAAAATAAAATTTCTTCAGGTTTTAAAA 1708

RESULT 4

US-09-899-651-5
; Sequence 5, Application US/09899651
; Patent No. 6576756
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/899,651
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/061,709

1 GCTGTAGCGAGGGGCTGGGGGCTGCTCTGCTCCCTTCTGCGGCTCGGCGCTCAGC 60
61 CCACCCAGAGGCGGGGTGGAGGGGAGTGTCTCAGCTTCCCGGGTTAGAGCGCGAAAA 120
61 CCACCCAGAGGCGGGGTGGAGGGGAGTGTCTCAGCTTCCCGGGTTAGAGCGCGAAAA 120
121 TTCAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTACTGGACAGCCTGCTGCTC 180
121 TTCAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTACTGGACAGCCTGCTGCTC 180
181 AGTATGTTAGTAGAGAACTGTGAGCAAGTGAACACCGAGAGTGAACCGCAGTGTGA 240
181 AGTATGTTAGTAGAGAACTGTGAGCAAGTGAACACCGAGAGTGAACCGCAGTGTGA 240
241 ATGTCACTATTCACCGGAGCAGACCCAGGCAAGCATCATGAAGCTGAATGCCACCC 300
241 ATGTCACTATTCACCGGAGCAGACCCAGGCAAGCATCATGAAGCTGAATGCCACCC 300
301 AGTTGGAGAACCATGCTGAGGTCTCTACATCCCGGATGAGCAGATAGCACAGGGAC 360
301 AGTTGGAGAACCATGCTGAGGTCTCTACATCCCGGATGAGCAGATAGCACAGGGAC 360
361 CTGAGAAATGGCGCGCGAGGGGCTTTGGCTCTCGGGTTCAGCCCGCGAGGGCTCACCTG 420
361 CTGAGAAATGGCGCGCGAGGGGCTTTGGCTCTCGGGTTCAGCCCGCGAGGGCTCACCTG 420
421 TGGCAGCGGGGCGCCAGCAGCAGCAAGTGAACATCCCGCTTCGGCTCTCTGCTGTC 480
421 TGGCAGCGGGGCGCCAGCAGCAGCAAGTGAACATCCCGCTTCGGCTCTCTGCTGTC 480
481 CCACCCAGTATGTGGGTGCCATTTATGGCAAGGAGGGGCGCACCATCCGCAATCAAA 540
481 CCACCCAGTATGTGGGTGCCATTTATGGCAAGGAGGGGCGCACCATCCGCAATCAAA 540
541 AACAGACCCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 600
541 AACAGACCCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 600
601 CCATCAGTGTGCTCCACCTCGAGGCTGCTCTCGCTTGTAAAGTGAATCTTGGAGA 660
601 CCATCAGTGTGCTCCACCTCGAGGCTGCTCTCGCTTGTAAAGTGAATCTTGGAGA 660
661 TTATGCATAAAGAGGCTTAAGACACAAACCGCTGACGAGGTTCCCTTGAAGTCTTGG 720
661 TTATGCATAAAGAGGCTTAAGACACAAACCGCTGACGAGGTTCCCTTGAAGTCTTGG 720
721 CCCATAAATACCTTTGTAGGGGCTCTCATTTGGCAAGGAGGAGCGAACCTTGAAGAGGTAG 780
721 CCCATAAATACCTTTGTAGGGGCTCTCATTTGGCAAGGAGGAGCGAACCTTGAAGAGGTAG 780
781 AGCAAGATACCGAGACAAATATCACCTCTCTGTTGCAAGACCTTACCTTTTACAACC 840
781 AGCAAGATACCGAGACAAATATCACCTCTCTGTTGCAAGACCTTACCTTTTACAACC 840
841 CTGAGAGGACCATCTGTTGAAGGGGCGCATCGAGAAATGTTTGAAGGGCGGAGGAGAAA 900
841 CTGAGAGGACCATCTGTTGAAGGGGCGCATCGAGAAATGTTTGAAGGGCGGAGGAGAAA 900
901 TAATGAAGAAATGTTGGAGGCGCTATGAGAAATGATGTGCTGCCATGAGCTCTCACCTGA 960
901 TAATGAAGAAATGTTGGAGGCGCTATGAGAAATGATGTGCTGCCATGAGCTCTCACCTGA 960
961 TCCCTGGCTGACCTGCTGCTGAGTCTTTTCCAGCTTCATCCAGCGAGTCCCGC 1020
961 TCCCTGGCTGACCTGCTGCTGAGTCTTTTCCAGCTTCATCCAGCGAGTCCCGC 1020
1021 CGCTCCAGCAGCGGTTACTGGGGCTGCTCCCTATAGTCTCTTTATGAGGCTCCGAGC 1080
1021 CGCTCCAGCAGCGGTTACTGGGGCTGCTCCCTATAGTCTCTTTATGAGGCTCCGAGC 1080
1081 AGAGATGTTGAGGTGTTTATCCCGGCGGAGGCTGGCGCATCATCGGCAAGAGG 1140
1081 AGAGATGTTGAGGTGTTTATCCCGGCGGAGGCTGGCGCATCATCGGCAAGAGG 1140

1141 GGACGACATCAACAGCTCTCCCGGTTTCCAGCGCTCCATCAAGATTCACACCCG 1200
1141 GGACGACATCAACAGCTCTCCCGGTTTCCAGCGCTCCATCAAGATTCACACCCG 1200
1201 AACACCTGACTCCAAAGTTTCTGATGTTATCATCTGACCGCCAGAGGCCCAATTCA 1260
1201 AACACCTGACTCCAAAGTTTCTGATGTTATCATCTGACCGCCAGAGGCCCAATTCA 1260
1261 AGGCTCAGGAAAGAAATCTATGGCAAACTCAAGGAGGAGAACTCTTTGGTCCCAAGGAG 1320
1261 AGGCTCAGGAAAGAAATCTATGGCAAACTCAAGGAGGAGAACTCTTTGGTCCCAAGGAG 1320
1321 AAGTGAAGCTGGAGACCCACATACGTCTGCCAGCATCAGCAGCTGGCGGCTCATTTGSCA 1380
1321 AAGTGAAGCTGGAGACCCACATACGTCTGCCAGCATCAGCAGCTGGCGGCTCATTTGSCA 1380
1381 AAGGTGAAACAAACCGGTGAAACGAGTTGAGAAATTTGACGGCAGCTGAGGTGTTAGTACAA 1440
1381 AAGGTGAAACAAACCGGTGAAACGAGTTGAGAAATTTGACGGCAGCTGAGGTGTTAGTACAA 1440
1441 GAGACAGACCCCTGATGAGAAACGACAGGTCATCGTGAATCATCGGACATTTCTATG 1500
1441 GAGACAGACCCCTGATGAGAAACGACAGGTCATCGTGAATCATCGGACATTTCTATG 1500
1501 CCAGTCAAGTGGCTCAACGGAGATCCGAGACATCTCGGCCAGCTTAAGCAGCAGCATC 1560
1501 CCAGTCAAGTGGCTCAACGGAGATCCGAGACATCTCGGCCAGCTTAAGCAGCAGCATC 1560
1561 AGAAGGACAGAGTAAACGAGGCCCGAGCAACGAGGAGTGAACAGCCCTCTCTGCTCCCT 1620
1561 AGAAGGACAGAGTAAACGAGGCCCGAGCAACGAGGAGTGAACAGCCCTCTCTGCTCCCT 1620
1621 TNGAGTCCAGACAAACCGGCGAGAAATCGAGAGTGTCTCTCCCGCAGGCTGAGA 1680
1621 TNGAGTCCAGACAAACCGGCGAGAAATCGAGAGTGTCTCTCCCGCAGGCTGAGA 1680
1681 ATGAGTGGGAATCCGGGACACNTGGCGGCTGTAGATCAGGTTTGGCCACTTGATTGA 1740
1681 ATGAGTGGGAATCCGGGACACNTGGCGGCTGTAGATCAGGTTTGGCCACTTGATTGA 1740
1741 GAAAGATGTTCCAGTGAAGAACCTTGATCTNTACGCCCCAAACACCCCAATTTGGGCC 1800
1741 GAAAGATGTTCCAGTGAAGAACCTTGATCTNTACGCCCCAAACACCCCAATTTGGGCC 1800
1801 AACACTGNTGCCCTCGGGGTGTAGAAATTTAGCCAGGCTTTTAAACCGTGGAT 1860
1801 AACACTGNTGCCCTCGGGGTGTAGAAATTTAGCCAGGCTTTTAAACCGTGGAT 1860
1861 TGTTTAAAGAGCTCTCCAGGCCCCACCAAGAGGTTGATCACCTCAGTGGGAGAAA 1920
1861 TGTTTAAAGAGCTCTCCAGGCCCCACCAAGAGGTTGATCACCTCAGTGGGAGAAA 1920
1921 AATAAAATTTCTTCAGGTTTAAAA 1946
1921 AATAAAATTTCTTCAGGTTTAAAA 1946

RESULT 3
US-09-061-709-5
; Sequence 5, Application US/09061709B
; Patent No. 6297364
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538

Db 301 AGTTGGACCAACATGTCCTGAGGTCCTCTACATCCCGATGAGCAGATAGCAGGGAC 360
QY 361 CTGAGATGGGGCGGAGGGGCTTTGGCTCTCGGGGTACGCCCGCAGGGCTCACCTG 420
Db 361 CTGAGATGGGGCGGAGGGGCTTTGGCTCTCGGGGTACGCCCGCAGGGCTCACCTG 420
QY 421 TGGCAGCGGGGGCCCCAGCCAGCAGCAGCAAGTGGACATCCCTTCCTGGCTCTGGTGC 480
Db 421 TGGCAGCGGGGGCCCCAGCCAGCAGCAGCAAGTGGACATCCCTTCCTGGCTCTGGTGC 480
QY 481 CCACCCAGTATGTGGGTGCCATTATTGGCAAGAGGGGGCCACCATCCGCAACATCAAA 540
Db 481 CCACCCAGTATGTGGGTGCCATTATTGGCAAGAGGGGGCCACCATCCGCAACATCAAA 540
QY 541 AACACAGCCAGTCCAAAGTAGACGTGCATAGGAGAGAGAACGCGAGTGCAGCTGAAAAG 600
Db 541 AACACAGCCAGTCCAAAGTAGACGTGCATAGGAGAGAGAACGCGAGTGCAGCTGAAAAG 600
QY 601 CCATCAGTGTGCACTCCACCCCTCAGGGCTGCTCCTCGCTTGTAAAGATGATCTGGAGA 660
Db 601 CCATCAGTGTGCACTCCACCCCTCAGGGCTGCTCCTCGCTTGTAAAGATGATCTGGAGA 660
QY 661 TTATGCATAAAGAGGCTTAAGACACCAAAACGGCTGACGAGGTTCCCTCGAGATCTGG 720
Db 661 TTATGCATAAAGAGGCTTAAGACACCAAAACGGCTGACGAGGTTCCCTCGAGATCTGG 720
QY 721 CCATATAACTTTGTAGGGGCTCTCAATGGCAAGGAGAACGCAACCTGAAGAAGTAG 780
Db 721 CCATATAACTTTGTAGGGGCTCTCAATGGCAAGGAGAACGCAACCTGAAGAAGTAG 780
QY 781 AGCAAGATACCGAGACAAAATCACCATCTCCTCGTTGCAAGACCTTACCCTTTTACAAC 840
Db 781 AGCAAGATACCGAGACAAAATCACCATCTCCTCGTTGCAAGACCTTACCCTTTTACAAC 840
QY 841 CTGAGAGGACCATCAGTGAAGGGGCCATCGAGATTTGTCAGGGCCGAGCAGGAAA 900
Db 841 CTGAGAGGACCATCAGTGAAGGGGCCATCGAGATTTGTCAGGGCCGAGCAGGAAA 900
QY 901 TAATGAAGAAAGTTCCGGAGGCTTATGAGATGATGCTGGCGCATGAGTCTCACCTGA 960
Db 901 TAATGAAGAAAGTTCCGGAGGCTTATGAGATGATGCTGGCGCATGAGTCTCACCTGA 960
QY 961 TCCTGGCTGAACCTGGCTGTAGTCTTTTCCAGCTTCATCCAGGCGAGTCCCGC 1020
Db 961 TCCTGGCTGAACCTGGCTGTAGTCTTTTCCAGCTTCATCCAGGCGAGTCCCGC 1020
QY 1021 CGCTCCAGCAGCTTACTGGGGCTGCTCCCTATAGCTCTTTATGAGGCTCCGAGC 1080
Db 1021 CGCTCCAGCAGCTTACTGGGGCTGCTCCCTATAGCTCTTTATGAGGCTCCGAGC 1080
QY 1081 AGGAGATGTCAGAGTGTATATCCCGCCAGCAGTGGCGCATCATCGCAAGAG 1140
Db 1081 AGGAGATGTCAGAGTGTATATCCCGCCAGCAGTGGCGCATCATCGCAAGAG 1140
QY 1141 GGCAGACATCAACAGCTCTCCGGTTTCCAGCGCTCCATCAAGATTCACACCCG 1200
Db 1141 GGCAGACATCAACAGCTCTCCGGTTTCCAGCGCTCCATCAAGATTCACACCCG 1200
QY 1201 ABAACCTGACTCCAAAGTTGATGGTTATCATCTGAGCGCGAGGGCCCAATCA 1260
Db 1201 ABAACCTGACTCCAAAGTTGATGGTTATCATCTGAGCGCGAGGGCCCAATCA 1260
QY 1261 AGGCTCAGGGAAGATCTATGCAAACTCAAGAGGAGAACTCTTGTGCCAGGAGG 1320
Db 1261 AGGCTCAGGGAAGATCTATGCAAACTCAAGAGGAGAACTCTTGTGCCAGGAGG 1320
QY 1321 AAGTGAAGCTGAGACCCACATAGGTGTGCCAGCATCAGCAGTGGCGGGTCATTGCA 1380
Db 1321 AAGTGAAGCTGAGACCCACATAGGTGTGCCAGCATCAGCAGTGGCGGGTCATTGCA 1380
QY 1381 AAGTGGAAAAACGGTGAACGATGTCAGATTTGACGCGACCTGAGTGGTAGTACCAA 1440

Db 1381 AAGTGGAAAAACGGTGAACGAGTTGCAGATTTTGCAGAGCTGAGTGGTAGTACCAA 1440
QY 1441 GAGACACAGACCCCTGATGAGAAACACAGGTCATCGTGAAATCATCGGACATTTCTATG 1500
Db 1441 GAGACACAGACCCCTGATGAGAAACACAGGTCATCGTGAAATCATCGGACATTTCTATG 1500
QY 1501 CCAGTCAAGTGGCTCAACGGAAGATCCGAGACATCTTGGCCAGGTTAAGCAGCAGATC 1560
Db 1501 CCAGTCAAGTGGCTCAACGGAAGATCCGAGACATCTTGGCCAGGTTAAGCAGCAGATC 1560
QY 1561 AGAAGGACACAGATTAACAGGCCACGAGCAGGAGGAAGTGAACAGCCCTTGTCCCT 1620
Db 1561 AGAAGGACACAGATTAACAGGCCACGAGCAGGAGGAAGTGAACAGCCCTTGTCCCT 1620
QY 1621 TNGAGTCCAGGACAAACACCGGGCAGAAATCGAGAGTGTCTCTCCCGGAGGCTCGAGA 1680
Db 1621 TNGAGTCCAGGACAAACACCGGGCAGAAATCGAGAGTGTCTCTCCCGGAGGCTCGAGA 1680
QY 1681 ATGAGTGGGAATCCGGGACACNTGGCGGGCTGTAGATCAGGTTTGGCCACTTGATGA 1740
Db 1681 ATGAGTGGGAATCCGGGACACNTGGCGGGCTGTAGATCAGGTTTGGCCACTTGATGA 1740
QY 1741 GAAAGATGTTCCAGTGAAGAACCTGATCTTTCAGCCCCAAAACACCCCAATTTGGGCC 1800
Db 1741 GAAAGATGTTCCAGTGAAGAACCTGATCTTTCAGCCCCAAAACACCCCAATTTGGGCC 1800
QY 1801 AACACTGTNTGCCCTCGGGTGTCTCAGAAATTTAGCGCAAGGCACCTTTTAAACGTGGAT 1860
Db 1801 AACACTGTNTGCCCTCGGGTGTCTCAGAAATTTAGCGCAAGGCACCTTTTAAACGTGGAT 1860
QY 1861 TGTTTAAAGAGTCTCCAGGCCCCACCAAGAGGGTGGATCAACCTCAGTGGGAAGAAA 1920
Db 1861 TGTTTAAAGAGTCTCCAGGCCCCACCAAGAGGGTGGATCAACCTCAGTGGGAAGAAA 1920
QY 1921 AATAAAATTCCTTCAGGTTTAAAA 1946
Db 1921 AATAAAATTCCTTCAGGTTTAAAA 1946

RESULT 2
US-09-899-651-7
; Sequence 7, Application US/09899651
; Patent No. 6576756
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; TITLE OF INVENTION: Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/899,651
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 7
; LENGTH: 1946
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-899-651-7

Query Match 99.7%; Score 1941; DB 4; Length 1946;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1946; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTGTAGCGAGGGCTGGGGGCTGCTCTCTCCCTTCTTGGCGCTGGGCTCAGC 60

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: July 24, 2004, 06:34:14 ; Search time 129.138 Seconds
(without alignments)
8362.646 Million cell updates/sec

Title: US-09-270-437D-7
Perfect score: 1946
Sequence: 1 gctgtacgaggggctggg.....attcttcaggttttaaaa 1946

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1941	99.7	1946	3	US-09-061-709-7
2	1941	99.7	1946	4	US-09-899-651-7
3	1669.2	85.8	1708	3	US-09-061-709-5
4	1669.2	85.8	1708	4	US-09-899-651-5
5	1301	66.9	2224	3	US-09-261-855-1
6	674.2	34.6	1740	4	US-09-643-597-347
7	674.2	34.6	1740	4	US-09-542-615A-347
8	674.2	34.6	1740	4	US-09-606-421B-347
9	672.6	34.6	4159	3	US-09-061-709-4
10	672.6	34.6	4159	4	US-09-899-651-4
11	672.6	34.6	4181	4	US-09-643-597-175
12	672.6	34.6	4181	4	US-09-480-884A-175
13	672.6	34.6	4181	4	US-09-542-615A-175
14	672.6	34.6	4181	4	US-09-606-421B-175
15	585.4	30.1	3283	3	US-09-061-709-8
16	585.4	30.1	3283	4	US-09-899-651-8
17	570.6	29.3	3412	3	US-09-061-709-6
18	570.6	29.3	3412	4	US-09-899-651-6
19	53.6	2.8	7218	1	US-08-232-463-14
20	45.6	2.3	598	4	US-09-669-751-39
21	40.6	2.1	2850	1	US-08-503-172-5
22	40.6	2.1	2850	2	US-09-135-211-5
23	39.6	2.0	1982	4	US-09-016-434-1067
24	39.6	2.0	1982	4	US-09-825-497A-40
25	39.6	2.0	1987	4	US-08-825-497A-39
26	37.6	2.0	1987	4	US-09-517-779-1
27	37.6	1.9	1896	4	US-09-343-011B-4

28	1.9	2608	4	US-09-154-750A-75	Sequence 75, Appl
29	1.9	2608	4	US-09-665-479A-11	Sequence 11, Appl
30	1.8	555	4	US-09-252-991A-6251	Sequence 6251, Ap
31	1.8	1644	4	US-09-252-991A-6420	Sequence 6420, Ap
32	1.8	2088	4	US-09-252-991A-6339	Sequence 6339, Ap
33	1.8	2202	4	US-09-252-991A-6131	Sequence 6131, Ap
34	1.8	2463	4	US-09-252-991A-6775	Sequence 6775, Ap
35	1.8	9862	4	US-09-691-861A-3	Sequence 3, Appli
36	1.8	765	4	US-09-252-991A-16348	Sequence 16348, A
37	1.8	867	4	US-09-252-991A-15936	Sequence 15936, A
38	1.8	505	4	US-09-621-976-15639	Sequence 15639, A
39	1.8	1386	2	US-08-910-731-5	Sequence 5, Appli
40	1.8	2517	4	US-09-894-998A-38	Sequence 38, Appl
41	1.8	2655	4	US-09-963-137-139	Sequence 139, App
42	1.8	2655	4	US-09-963-137-183	Sequence 183, App
43	1.8	832	4	US-09-621-976-2813	Sequence 2813, Ap
44	1.8	1926	4	US-09-249-585A-2	Sequence 2, Appli
45	1.8	1926	4	US-09-410-399-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-061-709-7
; Sequence 7, Application US/09061709B
; Patent No. 6297364
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/061,709B
; CURRENT FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 7
; LENGTH: 1946
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-061-709-7

Query Match	99.7%	Score 1941;	DB 3;	Length 1946;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1946;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	GCTGTAGCGAGGGGCTGGGGGGCTGCTCTGTCCTTCTTGGCGGCTGGCGGCTCAGC	60	
Db	1	GCTGTAGCGAGGGGCTGGGGGGCTGCTCTGTCCTTCTTGGCGGCTGGCGGCTCAGC	60	
QY	61	CCACCCAGAGCGGGGCTGGGGGGAGTGTCTACGCTTCCCGGTTAGAGCGGAGAAA	120	
Db	61	CCACCCAGAGCGGGGCTGGGGGGAGTGTCTACGCTTCCCGGTTAGAGCGGAGAAA	120	
QY	121	TTCAATCCGAATATTTCCACCCAGCTCCGATGGGAAGTACTGGACAGCTGCTGGCTC	180	
Db	121	TTCAATCCGAATATTTCCACCCAGCTCCGATGGGAAGTACTGGACAGCTGCTGGCTC	180	
QY	181	AGTATGGTACAGTAGAGAACTGTGAGCAAGTGAACCCAGAGTGAACCGGAGTGTGA	240	
Db	181	AGTATGGTACAGTAGAGAACTGTGAGCAAGTGAACCCAGAGTGTGAACCGGAGTGTGA	240	
QY	241	ATGTCACCTATTCCACCGGGAGCAGACGAGCAAGCATCATGAAGCTGAATGGCCACC	300	
Db	241	ATGTCACCTATTCCACCGGGAGCAGACGAGCAAGCATCATGAAGCTGAATGGCCACC	300	
QY	301	AGTTGGAGAACCATGCGCTGAAGGTTCTCTTACATCCCGGATGACAGATAGCAGGAC	360	

This Page Blank (uspto)

XX PF 28-JUN-2001; 2001WO-US021065.
XX PR 28-JUN-2000; 2000US-00606421.
XX PR 02-AUG-2000; 2000US-00630940.
XX PR 21-AUG-2000; 2000US-00643597.
XX PR 15-SEP-2000; 2000US-00662786.
XX PR 09-OCT-2000; 2000US-00685696.
XX PR 12-DEC-2000; 2000US-00735705.
XX PR 07-MAY-2001; 2001US-00850716.
XX (CORI-) CORIXA CORP.
XX Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
XX PI McNeill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;
XX PI Vedvick TS, Carter D, Watanabe Y, Peckham DW;
XX WPI; 2002-090513/12.
XX DR P-PSDB; ABB75053.
XX PT Polynucleotides encoding lung tumor polypeptides, useful for treating
XX PT lung cancer or stimulating an immune response.
XX PS Claim 1; Page 367; 374pp; English.
XX The present invention describes human lung tumour proteins. Human lung
XX CC tumour proteins and polynucleotides have cytostatic and immunostimulant
XX CC activities, and can be used in vaccine production. Compositions
XX CC comprising the lung tumour proteins, polynucleotides, antibodies, fusion
XX CC proteins, T cell populations, or antigen presenting cells that express
XX CC the lung tumour proteins are useful for treating lung cancer or
XX CC stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
XX CC ABB75070 represent sequences used in the exemplification of the present
XX CC invention
XX SQ Sequence 1743 BP; 527 A; 406 C; 418 G; 392 T; 0 U; 0 Other;
Query Match 34.6%; Score 674.2; DB 6; Length 1743;
Best Local Similarity 66.7%; Pred. No. 4.6e-177;
Matches 980; Conservative 0; Mismatches 483; Indels 6; Gaps 1;
QY 103 GGGTTAGGAGCGGGAATTCAGATCCGAAATATTCACCCAGCTCCGATCGGAGTAC 162
DB 230 GGCRAAGGATTCGGAATTCAGATACGAAATATTCGCCCTCATTTACAGTGGGAGTGC 289
QY 163 TGGCAGCCCTGCTGCTCAGTATGTCAGTAGAGAACTGTGAGCAAGTGAACACCGAGA 222
DB 290 TGGATAGTTTACTAGTCCAGTATGAGTGGTGGAGAGCTGTGAGCAAGTGAACATGACT 349
QY 223 GTGAGACGCGAGTGGTGAATGTCACTATTCCACCGGAGCAGACCGAGGCAAGCCATCA 282
DB 350 CGGAACCTGCACTGTAATGTAACTTATTCAGTAAAGGACCAAGCTAGACAAGCACTAG 409
QY 283 TGAAGCTGAATGGCCACAGTTGGAGAACCATATGCTCCTCAATCCCGGATG 342
DB 410 ACAAACTGAATGGATTCAGTTAGAGAAATTCACCTTGAAGTAGCCTATATCCCTGATG 469
QY 343 ACCAGATACCAAGGACCTGAGATGGCGCGGAGGGGCTTTGGCTCTCGGGGTGACG 402
DB 470 AAACGCGCCCGCAGCAAAACCCCTTGCAGACGCCCGAGGTGCGCGGGGGCTTGGGCAGA 529
QY 403 CCCGCGCAGGGCTCACCTGTGGCAGCGGGGGCCCCAGCCAGCAGCAGCAAGTGGACATCC 462
DB 530 GGGGCTCCTCAGGACGGGGTCTCCAGGATCCGATTCAGAGCAGAAACCATGTGATTTGC 589
QY 463 CCTTGGCTCTGTTGGCCACCCAGTATGCTGGTGGCCATTATTCGACAGGAGGGGCCA 522
DB 590 CTCTGGCTCTGTTGGTTCGCCCAATTTTGGAGGCCATCATAGGAAAGAGGTGCCA 649
QY 523 CCATCCGACATCACAAACAGACCCAGTCCAAAGATAGACGTGCTAGGAGGAGGAGAAC 582
DB 650 CCATTCCGAACATCACAAACAGACCCAGTCTAAATCGATGTCCACCGTAAAGAAATG 709

QY 583 CAGGTGACGCTGAAAAAGCCATCAGTGTGCACCTCCACCCCTGAGGGCTGCTCCTCCGCTT 642
DB 710 CGGGGCTGCTGAGAGTCGATTACTATCTCTCTACTCTCTGAGGCACTCTCTCGGCTT 769
QY 643 GTAAGATGATCTTTGGAGATTATCATAAAGAGCTAAGGACACCAAAACGGCTCAGCAGG 702
DB 770 GTAAGTCTATTCTGGAGATTATGCATAAGGAAGCTCAAGATATAAAATTCACAGAAGA 829
QY 703 TTCCCTGAAAGATCTTGGCCCATTAATAACTTTGTAGGGCTCTCATTTGGCAAGAGAC 762
DB 830 TCCCTTTGAAGATTTTAGCTCATATAACTTTGTGGACGCTCTATTGGTAAAGAGAA 889
QY 763 GGAACCTGAAAGAGGTAGAGCAAGATACCGAGACAAAATACCATCTCTCTGTTGCAAG 822
DB 890 GAATCTTTAAAAAAATTTAGCAGACACACACACTAAATACAGATATCTCCTATGCAAG 949
QY 823 ACCTTACCTTTTACACCTCTGAGAGGACCATCTACTGTGAAGGGGGCCCATCGAGAAATGTT 882
DB 950 AATTGACGCTGTATATCCAGAACGCACTATTACAGTTAAAGGCAATGTTGAGACATGTG 1009
QY 883 GCAGGGCCGAGCAGGAAATTAATGAGAAAGTTCCGGAGGCTATGAGAAATGATGTGGCTG 942
DB 1010 CCAAGCTGAGGAGGAGATCATGAAGAAAATCAGGGAGTCTTATGAAAATGATATGCTT 1069
QY 943 CCATGA-----GCTCTCAGCTGATCCCTGGCCCTGAACCTGGCTGTGTAGGTCTTTTCC 996
DB 1070 CTATGAATCTCAAGACATTTAATCTGGATTAATCTGAACGCTTGGGTCTGTGTCC 1129
QY 997 CAGCTTCCATCAGCGCAGTCCCGCCCTCCACAGCGTTTACTTGGGGCTCTCCCTATA 1056
DB 1130 CACCACTTCAGGGATGCCACTCCACCTCAGGGCCCTTCAGCCATGATCTCTCCCT 1189
QY 1057 GCTCTTTATGACGCTCCCGAGCAGGAGTGTGAGGTGTTTATCCCGCCCAAGGAG 1116
DB 1190 ACCCGCAGTTTGAGCAATTCAGAAACGGAGATGTTCTCTGTTTATCCAGCTCTATCAG 1249
QY 1117 TGGGCGCATCATCGCAAGAGGGGAGCAGCATCAACAGCTCTCCCGGTTTGCAGCG 1176
DB 1250 TCGGTGCCATCATCGCAGAGGCGGCGCAGCATCAACAGCTTCTCGCTTGTGGAG 1309
QY 1177 CTTCCATCAAGATTGCACACCCGAAACACCTGACTCCAAAGTTCGATGTTATGTTATCA 1236
DB 1310 CTTCAATTAAGATTGCTCCAGCGAAGCACCAGATGCTTAAAGTGAGGATGGTGAATTCA 1369
QY 1237 CTGAGCCCGCAGAGCCCAATTCAGGCTCAGGGAAGATCTATGCGAAACTCAAGGAGG 1296
DB 1370 CTGACACCAAGAGGCTCAGTTCAGGCTCAGGGGAAGATTTATGAAAAAATTAAGAG 1429
QY 1297 AGAATCTTTTGGTCCCAAGAGGAGGAAAGTGAAGCTGGAGACCCACATACGTGTCCAGCAT 1356
DB 1430 AAACTTTGTTAGTCTTAAAGAGAGGTGAACCTTGAAGCTCATATCAGAGTGCATCCT 1489
QY 1357 CAGCAGCTGGCGGGTCAATTTGGAAAAGCGTGAAGAGCGGTGAACGATTCGAGAAATTTGA 1416
DB 1490 TTGCTGCTGGCAGAGTTTATGGAAAAGGAGGCAAAACCGTGAATGAATTCAGAAATTTGT 1549
QY 1417 CGGAGCTGAGGTGTTAGTACCAAGAGCAGACCCCTGATGAGACGACGACGATCATCG 1476
DB 1550 CAAAGTGCAGAGTTTGTCTCTGTCGACACCTGATGAGATGATGAGATGACCAAGTGTG 1609
QY 1477 TGAAATCATCGACATTTCTATGCGAGTCAATGGCTCAACCGAAGATTCGAGACATCC 1536
DB 1610 TCAAAATTAAGTGTCTCTCTATGCTTGCAGTTGGCCAGAGTGGCCAGAGAAAATTCAGAAATTC 1669
QY 1537 TGGCCCGAGTTAAGCAGCAGCATCAGAG 1565
DB 1670 TGACTCAGGTAAGCAGCAGCAACCAACAG 1698

Search completed: July 24, 2004, 07:28:34
Job time : 725.632 secs

CC predetermined cut-off value, and determining the presence of a cancer in
CC the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
CC vaccines. (I) is useful as a marker to indicate the presence or absence
CC of a cancer such as lung cancer. AB092145 to AB092486 and ABP61866 to
CC ABP61992 represent sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 1740 BP; 526 A; 406 C; 417 G; 391 T; 0 U; 0 Other;

Query Match 34.68; Score 674.2; DB 6; Length 1740;
Best Local Similarity 66.74; Pred. No. 4.6e-177;
Matches 980; Conservative 0; Mismatches 483; Indels 6; Gaps 1;

QY 103 GGGTTAGGACCGGAAATTCAGATCGAAATATTCACCCAGCTCCGATGGGAATAC 162
DB 230 GCGAAGGATTCGGAATTCAGATCGAAATATTCACCCAGCTCCGATGGGAATAC 289

QY 163 TGGACAGCTGCTGGCTCAGTATGTTAGTACAGTACAGTACAGTACAGTACAGTAC 222
DB 290 TGGATAGTTTACTAGTCCAGTATGGAGTGGGAGAGCTGTGACAAAGTGAACACTGACT 349

QY 223 GTGAGCGGAGTGTGAATGTCACTTATCCACCGGAGCAGACAGCAAGCCATCA 282
DB 350 CGGAACCTGCGATGTTAAATGTAACTTATCCAGTAAGACCAAGCTAGACAGACTAG 409

QY 283 TGAAGCTGAATGGCCACAGTTGGAGAACCATGCTGAGGCTCTCTACATCCCGATG 342
DB 410 ACAACTGAATGATTTTCAGTACAGAAATTTCACTTGAAAGTAGCCTATATCCCTGATG 469

QY 343 AGCAGATAGCACAGGACCTGAGATGGGCGCCAGGGGGTCTGGCTCTCGGGGTGAGC 402
DB 470 AAACGGCGCCGACAAACCCCTTGCAGCAGCCCGAGGTCGCGGGGGTGGGGCAGA 529

QY 403 CCCGCCGGCTCCTCTGTGCGAGCGGGGGCCAGCCAGCAGCAGCAAGTGGACATCC 462
DB 530 GGGGCTCTCTAAGGAGGGGTCTCCAGGATCCGTATCCAGCAGAAACCATGTGATTTGC 589

QY 463 CCCTTCGGCTCCTGGTCCGACCCAGTATGTTGGTGCATTTATGGCAAGAGGGGGCCA 522
DB 590 CTCCTGGCTGCTGGTTCCTCCCAATTTGTTGGAGCCATCATAGGAAAGAGTGCCA 649

QY 523 CCATCCGCAACATCACAAAAGAGCCAGTCCAGATAGAGTGTGATAGGAGGAGAACG 582
DB 650 CCATTCGGAACATCACAAAGAGCCAGTCTAAATCGATGTCCACCCGTAAAGAAATG 709

QY 583 CAGGTGACGCTGAAAGGCCATCAGTGTGCACTCCACCCCTGAGGGCTCTCTCCGCTT 642
DB 710 CGGGGGCTGCTGAGAGTTCGATTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 769

QY 643 GTAAGATGATCTTGAGATTTATGCATAAAGAGGCTTAAGSACACCAAAACGGCTGACGAGG 702
DB 770 GTAAGTCTATTCTGGAGATTATGCATAAGGAAGCTCAAGATATAAAATTCACAGAAGAGA 829

QY 703 TTCCCTGAAGATCTGGCCCAATATACTTTGTAGGGGTCTCATTTGCAAGGAGGAC 762
DB 830 TCCCTTGAAGATTTTAGCTCATATACTTTGTTGGAGCTCTTATTTGTAAGAAGGAA 889

QY 763 GGAACCTGAGAGCTAGGAGAGATACCGACACAAAATCACTCTCTCTCTCTCTCTCT 822
DB 890 GAAATCTTAAAGAAATTTGAGCAAGACACAGACACTTAAATACAGATATCTCCATTCGAGG 949

QY 823 ACCTTACCTCTTACACCTCTGAGAGGACCATCACTGTGAAGGGGGCCATTCGAGAAATGTT 882
DB 950 AATTGAGCTGTATATCCAGACGACCTATTACAGTTAAAGCAATTTGAGACATGTG 1009

QY 883 GAGGCGCCAGCAGGAATTAATGAAGAGTTTCGGAGGCTTATGAGATGATGTTGGCTG 942
DB 1010 CCAAGCTCAGGAGGAGATCATGAAGAAATCAGGGAGTCTTATGAAATGATATTGCTT 1069

QY 943 CCATGA-----GCTCTACCTGATCCCTGGCTGAACTGGCTGCTGTAGTCTTTTCC 996
DB 1070 CTATGAATCTTCAGCACATTTAAATCTCTGGATTAATCTGAAGCCTTGGGCTCTGTCC 1129

QY 997 CAGCTTTCATCCAGCGAGTCCCGCGCTCCCGAGCAGCTTACTTGGGGCTGCTCCCTATA 1056
DB 1130 CACCACTTCAGGATGCGCCACTCCACCTCCAGCCCTTCCAGCCACTGACTCTCCCT 1189

QY 1057 GCTCTCTTTATGCAAGCTCCGAGCAGAGATGTTGAGGTTTATCCCGCCAGGAG 1116
DB 1190 ACCGCGAGTTTGAAGCAATCAGAAACGAGAGCTGTTCATCTGTTTATCCAGCTATCAG 1249

QY 1117 TGGGGCCATCATCGGCAAGAGGGGAGCAGACATCAACAGCTCTCCCGGTTTCCGACG 1176
DB 1250 TCGGTGCGATCATCGGCAAGCAGGCGCAGCAGCATCAAGCAGCTTCTCGCTTCTGAG 1309

QY 1177 CTTCCATCAAGATTGACACACCCGAAACACCTGACTCCAAAGTTGCTATGTTATCATCA 1236
DB 1310 CTTCAATTAAGATTGCTCCAGCGGAGCAACAGATGCTAAAGTGAGGATGCTGATTATCA 1369

QY 1237 CTGACCGCCAGAGGCCCAATTCAGGCTCAGGGAAGATCTATGGCAAACTCAAGGAG 1296
DB 1370 CTGACCAACAGAGGCTCAGTTCAAGGCTCAGGGAAGATTTATGGAATAATTAAGAAG 1429

QY 1297 AGAATCTTCTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGCTGCGCAGCAT 1356
DB 1430 AAAACTTTGTTAGTCTTAAGAGAGGAGTGAACCTTGAAGCTCATATCAGAGTGCATCT 1489

QY 1357 CAGCAGCTGGCGGGTCAATTGGCAAGGTGGAATAACCGTGAACAGAGTTGCAGAAATTGA 1416
DB 1490 TTGCTGCTGCGAGATTTATTGGAAAGAGGAGCAAAACGGTGAATGAACCTCAGAATTTGT 1549

QY 1417 CGGAGCTGAGGTGTTAGTACCAAGAGACGAGCCCTGTATGAGAACGACAGCTCATCG 1476
DB 1550 CAGTGCAGAAATTTGTTGCTCCTGTCACCAACACCTGTATGAGAAATACCAAGTGGTTG 1609

QY 1477 TGAATAATCATCGGACATTTCTATCCAGTCAAGTGGCTCAACGGGAAGATCCGAGACATCC 1536
DB 1610 TCAATAATACTGGTCACTTCTATGTTGCCAGTTGCCAGAGAAATTCAGGAATTC 1669

QY 1537 TGGCCCGAGTTAAAGCAGCAGCATCAGAAAG 1565
DB 1670 TGACTCAGTAAGAGCAGCACCACACACAG 1698

RESULT 14

ABL49299
ID ABL49299 standard; cDNA; 1743 BP.
XX
AC ABL49299;
XX
DT 01-MAY-2002 (first entry)
XX
DE Human lung tumour L523S recombinant polynucleotide SEQ ID NO:450.
XX
KW Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
XX
OS Homo sapiens.
XX
PN WO200200174-A2.
XX
PD 03-JAN-2002.
XX
PF 28-JUN-2001; 2001WO-US021065.
XX
PR 28-JUN-2000; 2000US-00606421.
XX
PR 02-AUG-2000; 2000US-00630940.
XX
PR 21-AUG-2000; 2000US-00643597.
XX
PR 15-SEP-2000; 2000US-00662786.
XX
PR 09-OCT-2000; 2000US-00685696.
XX
PR 12-DEC-2000; 2000US-00735705.
XX
PR 07-MAY-2001; 2001US-00850716.
XX
(CORI-) CORIXA CORP.
XX
PA Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
XX
PI

```
Db 290 TGGTAGTCTTACTAGTCCAGTATGAGTGTGGAGAGCTGTGAGCAAGTGAACACTGACT 349
Qy 223 GTGAGAGCGGCAAGTGGTGAATGTACACTATTCCAAACCGGAGCAGACCAAGGCAAGCCATCA 282
Db 350 CGGAAACTGCAAGTGTGAATGTAACTTATCCAGTAAGGACCAAGCTAGACAAGCACTAG 409
Qy 283 TGAAGCTCAATGGCCACCAAGTGTGAGAACCAATGCCCTGAAAGTCTCTACATCCCCCATG 342
Db 410 ACAAACCTGAATGATTTTCAGTTAGAGATTTTCACCTTGAAGTAGCCCTATATCCCTCATG 469
Qy 343 AGCAGATAGCACAGGACCTGAGAAATGGCGCCGAGGGGGCTTTGGCTCTCGGGGTGAGC 402
Db 470 AAACGGCGCGCCAGCAAAACCCCTTGACAGCAGCCCGAGGTGCGCGGGGGCTTGGGAGAG 529
Qy 403 CCGCCAGGGCTCACCTGTGGCAGCGGGGGCCCCAGCAAGCAGCAGCAAGGTGACATCC 462
Db 530 GGGGCTCCTCAAGCAGGGGTCTCCAGGATCCGATATCCAAAGCAGAAACCATGTGATTTGC 589
Qy 463 CCTTCGGCTCCTGGTGCCCAACCCAGTATGTGGGTGCAATTATTGGCAAGGAGGGGCCA 522
Db 590 CTCGTGCGCTGCTGGTTTCCCAOCCAAATTGTTGGAGCCATCATAGNAAGAAAGGTGCCA 649
Qy 523 CCATCCGCAATCATCAAAACAGACCCAGTCCAGATAGACGTGTCATAGGAAGGAGAACG 582
Db 650 CCATTCCGGAACATCACCAACAGACCCAGTCTAAATCGATGTCACCGTAAAGAAATG 709
Qy 583 CAGGTGCACTGAAAAAGCCATCAGTGTGCACCTCCACCCCTGAGGGTGTCTCTCGCTT 642
Db 710 CGGGGCTGCTGAGAGTCAATTAATCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 769
Qy 643 GTAAGATGATCTTTGGAGATTTATCATAAAGAGGCTTAAGGACACCAAAACGGCTGACGAG 702
Db 770 GTAAGTCTATTCTGGAGATTAATGATTAAGAACTCAAGATATAAATTCACAAAGAGAG 829
Qy 703 TTCCCTGAAGATCTCGGCCATTAATAAATTGTTAGGGGTCTCATTTGGCAAGGAGGAGAC 762
Db 830 TCCCTCTGAAGATTTTGTAGTCTAATAAATTGTTGGAGCTCTTATTGGTAAGAGGAA 889
Qy 763 GGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAATAATCAACATCTCTCTGTTGCAAG 822
Db 890 GAAATCTTAATAAATTTGAGCAGACACAGACACTAAATACGATATCTCAATTGACAG 949
Qy 823 ACCTTACCTTTACACCTTGAGAGACCATCACTGTGAAGGGGGCCATCGAGAAATTTT 882
Db 950 AATTGACGTGTATATCCAGAACGCATATTACAGTTAAGCAATGTTGAGACATGTG 1009
Qy 883 GCAGGCGCAGCAGGAAATTAAGAAAGTTCCGGAGGCCTATGAGAAATGATGGCTG 942
Db 1010 CCAAAGCTGAGGAGGAGATCATGAAGAAATCAGGGAGTCTTATGAAATGATATTGCTT 1069
Qy 943 CCATGA-----GCTCTCACCTGATCCCTGGCTGAACTGCTGTAGTCTTTTCC 996
Db 1070 CTATGAATCTTCAAGCACATTAATTTCTGGATTAATCTGAAAGCTTTGGGTCTGTTC 1129
Qy 997 CAGTCTATCCAGCGAGTCCGCGCCCTCCAGCAGCGTTACTGGGGCTGCTCCCTATA 1056
Db 1130 CACCCACTTCAGGATGCCACCTCCACCTCAGGCGCCCTTCAGCCATGACTCTCCCT 1189
Qy 1057 GCTCTTTATGAGGCTCCGAGCAGGATGCTGAGGTGTTTATCCCGCCAGGAG 1116
Db 1190 ACCCGAGTTTGAGCAATCAGAAACGGAGACTGTTTCATCTGTTTATCCAGCTCATCAG 1249
Qy 1117 TGGGCGCCATCATCGCAAGAGGGGCGACATCAACAGCTCTCCCGGTTTGGCAGG 1176
Db 1250 TCGGTGCCATCATCGCAAGAGGGGCCAGACATCAAGCAGTTTCTCGCTTGTGAG 1309
Qy 1177 CTTCAATTAAGATTGCTCCAGCGAAGCACCAGATGCTTAAGTGAAGTGTGATATCA 1236
Db 1310 CTTCAATTAAGATTGCTCCAGCGAAGCACCAGATGCTTAAGTGAAGTGTGATATCA 1369
Qy 1237 CTGAGCCCGAGGCCCAATTCAAGGCTCAGGGAAGATCTATGGCAAACTCAAGGAG 1296
```

```
Db 1370 CTGACCAACCAAGGCTCAGTTCAAGGCTCAGGAGAGATTTATGGAATAATTAAAGAA 1429
Qy 1297 AGAACTTTCTTTTGTGCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTTGTCACGAT 1356
Db 1430 AAACTTTGTTAGTCTTAAGAGAGAGTGAACCTTGAAGCTCATATCAGAGTGCCTCCT 1489
Qy 1357 CAGCAGTGCAGGCTCATTTGGCAAGGTGGAATAAAACGGTGAACGAGTTTGCAGAAATTGA 1416
Db 1490 TTGCTGTGTCAGAGTTATTGGAAAGAGGAGCAAAACGGTGAATGAACCTTCAAGATTGT 1549
Qy 1417 CGCAGCTGAGTGGTAGTACCAAGAGACCAAGCCCTGATGAGACGACCAAGGTCTATCG 1476
Db 1550 CAAGTGCAGAAAGTTGTTGTCCTCGTGACCAAGACACTGATGAGAAATTCAGAAATTC 1609
Qy 1477 TGAATAATCATCGACATTTCTATGCCAGTCAAGTGGTCAACGGAAGATCCGAGACATCC 1536
Db 1610 TCAAAATACTGCTCACTTCTATGCTTGCAGGTTGCCAGAGAAATTCAGAAATTC 1669
Qy 1537 TGSCCCAGGTTAAGCAGCAGCATCAGAA 1565
Db 1670 TGACTCAGGTAAGCAGCACCACAACAG 1698

RESULT 13
ABQ92440
ID ABQ92440 standard; cDNA; 1740 BP.
XX
AC ABQ92440;
XX
DT 07-OCT-2002 (first entry)
XX
DE Human lung cancer associated cDNA sequence SEQ ID NO:347.
XX
KW Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine; gene;
XX ss.
XX Homo sapiens.
XX
PN WO200247534-A2.
XX
PD 20-JUN-2002.
XX
PF 30-NOV-2001; 2001WO-US047576.
XX
PR 12-DEC-2000; 2000US-00735705.
PR 07-MAY-2001; 2001US-00850716.
PR 28-JUN-2001; 2001US-00897778.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX
DR WPI; 2002-583465/62.
XX
P-PSDB; ABP61917.
XX
PT Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
PT the polynucleotides, useful in pharmaceutical compositions such as
PT vaccines and as markers to indicate the presence of lung cancer.
XX
PS Claim 1; Page 337; 381pp; English.
XX
CC The present invention describes isolated human lung carcinoma
CC polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
CC activity, and can be used in gene therapy and in vaccines. Compositions
CC comprising (I) or (II) can be used for stimulating an immune response in
CC a patient and for treating lung cancer in a patient. Oligonucleotides of
CC (I) can be used for detecting the presence of a cancer in a patient, by
CC obtaining a biological sample from the patient, contacting the biological
CC sample with the oligonucleotide, detecting in the sample, an amount of
CC polynucleotide that hybridises to the oligonucleotide and comparing the
CC amount of polynucleotide that hybridises to the oligonucleotide to a
```


CC methods. The present sequence was used as the representative sequence
CC from a human clone which was used in homology searches to identify the
CC clone. Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in CD-ROM format directly from
CC EPO
XX
SQ Sequence 833 BP; 223 A; 225 C; 215 G; 166 T; 0 U; 4 Other;
Query Match 37.7%; Score 732.8; DB 4; Length 833;
Best Local Similarity 96.9%; Pred. No. 1.5e-193;
Matches 802; Conservative 0; Mismatches 14; Indels 12; Gaps 5;
QY 537 ACAAACAGACCCAGTCCAGATAGACGTGCATAGGAGAGAACGCGAGTGCAGCTGAA 596
Db 1 ACAAACAGACCCAGTCCAGATAGACGTGCATAGGAGAGAACGCGAGTGCAGCTGAA 60
QY 597 AAAGCCATCAGTGTGACTCCACCCCTGAGGCTGCTCCTCCGCTTGAAGATGATCTTG 656
Db 61 AAAGCCATCAGTGTGACTCCACCCCTGAGGCTGCTCCTCCGCTTGAAGATGATCTTG 120
QY 657 GAGATTATGATAAAGAGGCTAAGGACACCAAAACGCTGACGAGTTCCCTGAGATC 716
Db 121 GAGATTATGATAAAGAGGCTAAGGACACCAAAACGCTGACGAGTTCCCTGAGATC 180
QY 717 CTGGCCCATTAATACTTTGTAGGCGCTCATTTGGCAAGAGGACGGAACTTGAAGAAG 776
Db 181 CTGGCCCATTAATACTTTGTAGGCGCTCATTTGGCAAGAGGACGGAACTTGAAGAAG 240
QY 777 GTAGACAGATACCCGAGACAAATAATCAATCTCTCGTTGCAAGACCTTACCTTTAC 836
Db 241 GTAGACAGATACCCGAGACAAATAATCAATCTCTCGTTGCAAGACCTTACCTTTAC 300
QY 837 AACCTCTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAAATTTGCGAGGCGGAGCAG 896
Db 301 AACCTCTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAAATTTGCGAGGCGGAGCAG 360
QY 897 GAAATTAATGAAGAAATTCGGAGGCGCTATGAGAAATGATGTGCTGCCATGAGC----- 950
Db 361 GAAATTAATGAAGAAATTCGGAGGCGCTATGAGAAATGATGTGCTGCCATGAGCCTGCAG 420
QY 951 TCTACCTCATCCCTGGCTGACCTGACCTGGCTGCTGTAGGTCTTTTCCAGCTTCATCCAGC 1010
Db 421 TCTACCTCATCCCTGGCTGACCTGACCTGGCTGCTGTAGGTCTTTTCCAGCTTCATCCAGC 480
QY 1011 GCAGTCCCGCGCTCCCGAGCGCTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCGAG 1070
Db 481 GCAGTCCCGCGCTCCCGAGCGCTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCGAG 540
QY 1071 GCTCCCGAGCAGAGATGTGAGTGTATATCCCGCCGAGGCGAGTGGCGGCCCATCATC 1130
Db 541 GCTCCCGAGCAGAGATGTGAGTGTATATCCCGCCGAGGCGAGTGGCGGCCCATCATC 600
QY 1131 GGCAAGAGGGGCGAGCAGATCAAAAGAGCTCTCCGGTTTGGCAGCGCTTCCATCAAGATT 1190
Db 601 GGCAAGAGGGGCGAGCAGATCAAAAGAGCTCTCCGGTTTGGCAGCGCTTCCATCAAGATT 660
QY 1191 GCACACCCGAAACCTGACTCCAAAGTTCCGTATGGTTATCATCACTGGA-CCGCCAGA 1249
Db 661 GCACACCCGAAACCTGACTCCAAAGTTCCGTATGGTTATCATCACTGGA-CCGCCAGA 720
QY 1250 GGCCCAATTCAGGCTCAGGGAAGAAATCATGGCAAACTC-AAGGAGGAGAACTT--CTT 1306
Db 721 NGCCCAATTCAGGCTCAGGGAAGAAATCATGGCAAACTC-AAGGAGGAGAACTTCTT 780
QY 1307 TGGTCCCAAGGAGGAGT--GAGAGTGGAGACCCACATACGTGTGCCA 1352
Db 781 TGGTCCCAAGGAGGAGT--GAGAGTGGAGACCCACATACGTGTGCCA 828
RESULT 11
AAC66035
ID AAC66035 standard, cDNA; 1740 BP.
XX

AAC66035;
XX 21-FEB-2001 (first entry)
XX Human lung cancer-associated cDNA antigen L523S.
XX
XX Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;
XX vaccine; detection; ss.
XX Homo sapiens.
XX WO200061612-A2.
XX 19-OCT-2000.
XX 03-APR-2000; 2000WO-US008896.
XX 02-APR-1999; 99US-00285479.
XX 17-DEC-1999; 99US-00466396.
XX 30-DEC-1999; 99US-00476496.
XX 10-JAN-2000; 2000US-00480884.
XX 22-FEB-2000; 2000US-00510376.
XX (CORI-) CORIXA CORP.
XX Wang T, Fan L;
XX WPI; 2000-628399/60.
XX P-PSDB; AAB11365.
XX Isolated polypeptide comprising an immunogenic portion of a lung tumor
XX protein is used for detecting and monitoring progression of lung cancer
XX in a patient.
XX Claim 1a; Page 258-259; 261pp; English.
XX This invention describes a novel isolated polypeptide (I) which
XX comprising an immunogenic portion of a lung tumor protein or variant (P2)
XX which have cytostatic activity. The polypeptides and polynucleotides are
XX used in compositions and vaccines to inhibit the development of cancer,
XX especially lung cancer, in a patient. Methods described in the invention
XX can be used to monitor the progression of a cancer by carrying out the
XX detection at subsequent time points and comparing the results from the
XX different time points. CD4+ and/or CD8+ T-Cells isolated from a patient
XX are treated with P2, polynucleotides encoding P2 or antigen presenting
XX cells expressing P2 and then administered to the patient to inhibit
XX development of cancer
XX
SQ Sequence 1740 BP; 526 A; 406 C; 417 G; 391 T; 0 U; 0 Other;
Query Match 34.6%; Score 674.2; DB 3; Length 1740;
Best Local Similarity 66.7%; Pred. No. 4.6e-177;
Matches 980; Conservative 0; Mismatches 483; Indels 6; Gaps 1;
QY 103 GGGTTAGAGCGCGGAAATTCAAATCCCAATATTTCCACCCAGCTCCGATGGGAGTAC 162
Db 230 GCGAAGGATTCGGAACCTTCAGATACGAAATATCCCGCTCAATTTACAGTGGGAGTGC 289
QY 163 TGGACAGCCTCGTGGCTCAGTATGGTACAGTAGAAGTGTGAGCACTGAGCACTGACCGAGA 222
Db 290 TGGATAGTTTACTAGTCAGTATGGAGTGGTGGAGCTGTGAGCACTGAGCACTGACT 349
QY 223 GTGAGACGCGAGTGGTGAATCTACCTATTCCAAACCGGAGCAGCAGGCAAGCAATCA 282
Db 350 CGGAACTGCAAGTTGTAATGTAACCTATTCCAGTAAGGACCAAGCTAGACAAGCACTAG 409
QY 283 TGAAGCTGAATGGCCACAGTGGAGAACCAATGCCCTGAAGTCTCTCATATCCCGGATG 342
Db 410 ACAACTGAATGGATTTTCAGTTAGAGAAATTTTCAGCTTGAAGTAGCTTATATCCCTGATG 469
QY 343 AGCAGATAGCACAGGACCTGAGAAATGGGCGCCGAGGGGGCTTTGGCTCTCGGGGTGAGC 402
Db 470 AACGGCGCCCGACCAAAACCCCTTGACAGAGCCCGAGGTGCGCGGGGGCTTGGGCGAGA 529

KW	Human; full length cDNA; cDNA synthesis; oligo-capping; ss.	951	TCTCACTGATCCCTGGCCTGAACCTGCTGCTAGGTCTTTTCCAGCTTCATCCAGC	1010
XX				
OS	Homo sapiens.	421	TCTCACTGATCCCTGGCCTGAACCTGCTGCTAGGTCTTTTCCAGCTTCATCCAGC	480
XX				
PN	EP1130094-A2.	1011	GCAGTCCCCTGGCCTCCAGCAGCGTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCAG	1070
XX				
PD	05-SEP-2001.	481	GCAGTCCCCTGGCCTCCAGCAGCGTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCAG	540
XX				
PF	07-JUL-2000; 2000EP-00114089.	1071	GCTCCGAGCAGAGATGCTGAGGTGTTTATCCCGCCAGCAGTGGCGGCATCATC	1130
XX				
XX	08-JUL-1999; 99JP-00194486.	541	GCTCCGAGCAGAGATGCTGAGGTGTTTATCCCGCCAGCAGTGGCGGCATCATC	600
PR	11-JAN-2000; 2000JP-00118774.	1131	GGCAAGAAGGGGAGCAGACATCAACAGCTCTCCCGTTTCCAGCGCTCCATCAAGATT	1190
PR	02-MAY-2000; 2000JP-00183765.	601	GGCAAGAAGGGGAGCAGACATCAACAGCTCTCCCGTTTCCAGCGCTCCATCAAGATT	660
XX				
XX	(HELI-) HELIX RES INST.	1191	GCACACCCGAAACACTGACTCCAAAGTTGCTATGTTATCATCATCTGA-CGCGCAGA	1249
XX	Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;			
PI	Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;	661	GCACACCCGAAACACTGACTCCAAAGTTGCTATGTTATCATCATCTGA-CGCGCAGA	720
XX				
DR	WPI; 2001-524255/58.	1250	GGCCCAATTCAGGCTCAGGGAAGATCTATGGCAATC-AAGGAGGAGAACTT--CTT	1306
XX				
XX	830 Primers useful for synthesizing full length cDNA clones and their use	721	NGCCCAATTCAGGCTCAGGGAAGATCTATGGCAATC-AAGGAGGAGAACTT--CTT	780
PT	in genetic manipulation.			
XX		1307	TGCTCCCAAGCAGGAAGT--GAAGCTGGAGACCCACATACGTGTGCCA	1352
PS	Claim 2; SEQ ID NO 429; 1380pp + Sequence Listing; English.			
XX		781	GGTCCCAAGGAGGAAGTTGAAGCTTGGAGACCCACATACCTGTGCCA	828
XX				
CC	The invention relates to primers for synthesizing full length cDNA			
CC	clones. 830 cDNA molecules encoding a human protein have been isolated			
CC	and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have			
CC	been determined. Primers for synthesizing the full length cDNA are useful			
CC	for clarifying the function of the protein encoded by the cDNA. The full			
CC	length clones were obtained by construction of full length enriched cDNA			
CC	libraries that were synthesised by the oligo-capping method. The primers			
CC	enable the production of the full length cDNA easily without any special			
CC	methods. The present sequence is the nucleotide sequence of the 5'-end of			
CC	a cDNA provided in the invention. Note: The sequence data for this patent			
CC	did not form part of the printed specification, but was obtained in CD-			
CC	ROM format directly from EPO			
XX				
XX	Sequence 833 BP; 223 A; 225 C; 215 G; 166 T; 0 U; 4 Other;			
SQ				
	Query Match 37.7%; Score 732.8; DB 4; Length 833;			
	Best Local Similarity 96.9%; Pred. NO. 1.5e-193;			
	Matches 802; Conservative 0; Mismatches 14; Indels 12; Gaps 5;			
Qy	537 ACAAAACAGACCCAGTCCAGATAGACGTGATAGAGAGAGACGAGGTGAGCTGAA	596		
Db	1 ACAAAACAGACCCAGTCCAGATAGACGTGATAGAGAGAGACGAGGTGAGCTGAA	60		
Qy	597 AAAGCCATCAGTGTGCTCACTCCACCCCTGAGGGCTGCTCTCCGCTTTGTAAGATGCTTG	656		
Db	61 AAAGCCATCAGTGTGCTCACTCCACCCCTGAGGGCTGCTCTCCGCTTTGTAAGATGCTTG	120		
Qy	657 GAGATTATGCAATAAGAGGCTAAGACACCAAAACGGCTGACAGGTTCCCTGAGATC	716		
Db	121 GAGATTATGCAATAAGAGGCTAAGACACCAAAACGGCTGACAGGTTCCCTGAGATC	180		
Qy	717 CTGGCCCATATAACTTTGTAGGGCTCTCATTTGCAAGGAAGGACGGAACCTGAAGAAG	776		
Db	181 CTGGCCCATATAACTTTGTAGGGCTCTCATTTGCAAGGAAGGACGGAACCTGAAGAAG	240		
Qy	777 GTAGAGCAAGATACCGAGACAAATAATCACCATCTCTCTGTCAGACCTTACCCCTTAC	836		
Db	241 GTAGAGCAAGATACCGAGACAAATAATCACCATCTCTCTGTCAGACCTTACCCCTTAC	300		
Qy	837 AACCTGTAGAGACCATCATCTGTGAAGGGGCCATCAGAAATTTGTCAGGGCCGAGCAG	896		
Db	301 AACCTGTAGAGACCATCATCTGTGAAGGGGCCATCAGAAATTTGTCAGGGCCGAGCAG	360		
Qy	897 GAAATATGAGAAAGTTCGGAGGGCTTATGAAATGATGTGGCTGCCATGAGC-----	950		
Db	361 GAAATATGAGAAAGTTCGGAGGGCTTATGAAATGATGTGGCTGCCATGAGCCTGCAG	420		

CC The invention relates to primers for synthesizing full length cDNA clones. 830 cDNA molecules encoding a human protein have been isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have been determined. Primers for synthesizing the full length cDNA are useful for clarifying the function of the protein encoded by the cDNA. The full length clones were obtained by construction of full length enriched cDNA libraries that were synthesised by the oligo-capping method. The primers enable the production of the full length cDNA easily without any special

27-SEP-2000; 2000US-0235834P.
 29-SEP-2000; 2000US-0236327P.
 29-SEP-2000; 2000US-0236367P.
 29-SEP-2000; 2000US-0236368P.
 29-SEP-2000; 2000US-0236369P.
 29-SEP-2000; 2000US-0236370P.
 02-OCT-2000; 2000US-0236802P.
 02-OCT-2000; 2000US-0237037P.
 02-OCT-2000; 2000US-0237038P.
 02-OCT-2000; 2000US-0237039P.
 13-OCT-2000; 2000US-0237040P.
 20-OCT-2000; 2000US-0239935P.
 20-OCT-2000; 2000US-0240960P.
 20-OCT-2000; 2000US-0241785P.
 01-NOV-2000; 2000US-0241809P.
 17-NOV-2000; 2000US-0244617P.
 08-DEC-2000; 2000US-0249299P.
 08-DEC-2000; 2000US-0251856P.
 08-DEC-2000; 2000US-0251868P.
 08-DEC-2000; 2000US-0251869P.
 (ROSE/) ROSEN C A.
 (RUBE/) RUBEN S M.
 (BARA/) BARASH S C.
 Rosen CA, Ruben SM, Barash SC;
 WPI; 2003-147444/14.
 P-PSDB; ABUS5647.
 New polypeptides and nucleic acids, useful in gene therapy for treating,
 inhibiting or preventing e.g. neural, immune system, muscular,
 respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
 renal disorders.
 Claim 1; SEQ ID NO 745; 402pp; English.
 The invention relates to human novel polypeptides and their associated
 polynucleotides. The polypeptides and polynucleotides are useful in gene
 therapy for treating, inhibiting or preventing neural disorders, immune
 system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
 and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
 nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
 gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
 (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
 heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
 renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
 leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
 appendicitis), allergic reactions and conditions (e.g. asthma), blood
 related disorders (e.g. thrombosis, atherosclerosis and myocardial
 infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent
 human novel polynucleotides of the invention
 Query Match 55.0%; Score 1069.8; DB 7; Length 1129;
 Best Local Similarity 98.7%; Pred. No. 2e-287;
 Matches 1084; Conservative 3; Mismatches 5; Indels 6; Gaps 1;
 QY 855 ACTGTGAAGGGGGCCATCGAATTTGTCAGGGCCGAGCGAGGAATAATGAAGAAAGTT 914
 Db 1 ACTGTGAAGGGGGCCATCGAATTTGTCAGGGCCGAGCGAGGAATAATGAAGAAAGTT 60
 QY 915 CGGAGGGCCATGAGATGATGTGGCTGCCATGAGC-----TCTCAGCTGATCCCTGGC 968
 Db 61 CGGAGGGCCATGAGATGATGTGGCTGCCATGAGCCTGCGAGCTTCACCTGATCCCTGGC 120
 QY 969 CTGAACCTGGCTGCTGTAGGTCCTTTTCCAGCTTCATCCAGCGAGTCGCCCGCCTCCC 1028
 Db 121 CTGAACCTGGCTGCTGTAGGTCCTTTTCCAGCTTCATCCAGCGAGTCGCCCGCCTCCC 180
 QY 1029 AGCAGCGTTACTGGGGCTGCTCCCTTAGCTTCCTTTATGCAAGGCTCCGAGCAGAGATG 1088

Db 181 AGCAGCGTTACTGGGGCTGCTCCCTATAGCTCCTTTATGAGGCTCCCGAGCAGGAGATG 240
 QY 1089 GTGCAGGTGTTTATCCCGCCAGGCGAGTGGGCCCATCATCGCAGAGAGGGGCGAGAC 1148
 Db 241 GTGCAGGTGTTTATCCCGCCAGGCGAGTGGGCCCATCATCGCAGAGAGGGGCGAGAC 300
 QY 1149 ATCAACAGCTCTCCCGGTTTGCAGCGCTCCATCAAGATTGCACCAACCCGAAACACCT 1208
 Db 301 ATCAACAGCTCTCCCGGTTTGCAGCGCTCCATCAAGATTGCACCAACCCGAAACACCT 360
 QY 1209 GACTCCAAAGTTGTTATGTTATCATCTGGAACCGCCAGAGGCCCAATTAAGGCTCAG 1268
 Db 361 GACTCCAAAGTTGTTATGTTATCATCTGGAACCGCCAGAGGCCCAATTAAGGCTCAG 420
 QY 1269 GGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTTCTTTGGTCCCGAGGAGGAGTAAG 1328
 Db 421 GGAAGAATCTATGCGAATCAAGGAGGAGAACTTTCTTTGGTCCCGAGGAGGAGTAAG 480
 QY 1329 CTGGAGACCCACATACGTGTGCGAGCATCAGCAGCTGGCGGGTCAATTGGCAAGGTGGA 1388
 Db 481 CTGGAGACCCACATACGTGTGCGAGCATCAGCAGCTGGCGGGTCAATTGGCAAGGTGGA 540
 QY 1389 AAAACGGTGAACGAGTTGCGAATTTGACGGGAGCTGAGGTGAGTAGTACCAAGAGACCG 1448
 Db 541 AAAACGGTGAACGAGTTGCGAATTTGACGGGAGCTGAGGTGAGTAGTACCAAGAGACCG 600
 QY 1449 ACCCTGTATGAGAACGACGAGGTTCATCGTGAATAATCATCGGACATTTCTATGCCAGTCAG 1508
 Db 601 ACCCTGTATGAGAACGACGAGGTTCATCGTGAATAATCATCGGACATTTCTATGCCAGTCAG 660
 QY 1509 ATGGCTCAACGAAAGATCCAGACATCTCTGGCCAGGTTAAAGCAGCAGCATCAGAAAGGA 1568
 Db 661 ATGGCTCAACGAAAGATCCAGACATCTCTGGCCAGGTTAAAGCAGCAGCATCAGAAAGGA 720
 QY 1569 CAGAGTTACCGAGCCCGCAGCAGGAGGAGTAGCAGAGCCCTCCCTGTCCCTTNGAGTCC 1628
 Db 721 CAGAGTTACCGAGCCCGCAGCAGGAGGAGTAGCAGAGCCCTCCCTGTCCCTTNGAGTCC 780
 QY 1629 AGGACAAACAAACGGGCGAGAAATCGAGAGTGTGCTCTCCCGGCGAGGCTGAGAAATGAGTGG 1688
 Db 781 AGGACAAACAAACGGGCGAGAAATCGAGAGTGTGCTCTCCCGGCGAGGCTGAGAAATGAGTGG 840
 QY 1689 GAATCCGGGACACNTGGCGCGGGCTGTAGATCAGGTTTGCACATTTGATTGAGAAAGATG 1748
 Db 841 GAATCCGGGACACNTGGCGCGGGCTGTAGATCAGGTTTGCACATTTGATTGAGAAAGATG 900
 QY 1749 TTCCAGTCAGGAACCTCATCTNTCAGCCCAACCAACACCCCAATTTGCCCAACACTGT 1808
 Db 901 TTCCAGTCAGGAACCTCATCTNTCAGCCCAACCAACACCCCAATTTGCCCAACACTGT 1808
 QY 1809 NTGCCCTCCGGGGTGTGAGAAATTTAGCGCAAGGCACTTTTAAACGTGGATTGTTTAA 1868
 Db 961 NTGCCCTCCGGGGTGTGAGAAATTTAGCGCAAGGCACTTTTAAACGTGGATTGTTTAA 1020
 QY 1869 GAAGCTCTCCAGGCGCCACCAAGAGGGTGGATCACTCAGTGGGAGAAAAATAAAT 1928
 Db 1021 GAAGCTCTCCAGGCGCCACCAAGAGGGTGGATCACTCAGTGGGAGAAAAATAAAT 1080
 QY 1929 TTCCTTCAGGTTTAAAA 1946
 Db 1081 TTCCTTCAGGTTTAAAA 1098
 RESULT 9
 AAK91969
 ID AAK91969 standard; cDNA; 833 BP.
 XX AAK91969;
 AC AAK91969;
 XX
 DT 06-NOV-2001 (first entry)
 XX Human cDNA 5'-end sequence, SEQ ID NO: 429.
 DE
 XX

CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence encodes a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Query Match		55.0%;	Score 1069.8;	DB 4;	Length 1129;
Best Local Similarity		98.7%;	Pred. No. 2e-287;	5;	Gaps 1;
Matches 1084;		Conservative	3;	Mismatches	5;
QY	855	ACTGTCAAGGGGCCATCGAATATGTTGCGAGGGCCGAGCAGCAATAATGAAGAAGTT	914		
DB	1	ACTGTGAAGGGGCCATCGAATATGTTGCGAGGGCCGAGCAGCAATAATGAAGAAGTT	60		
QY	915	CGGAGGCGCTATGAGAAATGATGTGGCTGCATGAGC-----TCTACCTGATCCCTGGC	968		
DB	61	CGGAGGCGCTATGAGAAATGATGTGGCTGCATGAGCCTGAGCTCTCACCTGATCCCTGGC	120		
QY	969	CTGAACCTGCTGCTGTAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCGCTCC	1028		
DB	121	CTGAACCTGCTGCTGTAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCGCTCC	180		
QY	1029	AGCAGGCTTACTCGGGCTGCTCCCTATAGTCTCTTTATGAGGCTCCCGAGCAGGATG	1088		
DB	181	AGCAGGCTTACTCGGGCTGCTCCCTATAGTCTCTTTATGAGGCTCCCGAGCAGGATG	240		
QY	1089	GTGAGGCTGTTATCTCCCGCCAGGCGAGTGGCGCCATCATCGCAGAGAGGGCAGCAC	1148		
DB	241	GTGAGGCTGTTATCTCCCGCCAGGCGAGTGGCGCCATCATCGCAGAGAGGGCAGCAC	300		
QY	1149	ATCAAAAGCTCTCCGGTTGCGAGCGCTCCATCAAGTTGACACCCGAAACACCT	1208		
DB	301	ATCAAAAGCTCTCCGGTTGCGAGCGCTCCATCAAGTTGACACCCGAAACACCT	360		
QY	1209	GACTCCAAAGTTCGTATGGTATCATCACTGAGCGCGCAGAGGCGCCAAATCAAGGCTCAG	1268		
DB	361	GACTCCAAAGTTCGTATGGTATCATCACTGAGCGCGCAGAGGCGCCAAATCAAGGCTCAG	420		
QY	1269	GGAGAATCTATGGCAACTCAAGGAGGAGAACTTTTGTGTCGCAAGGAGGAGTGAAG	1328		
DB	421	GGAGAATCTATGGCAACTCAAGGAGGAGAACTTTTGTGTCGCAAGGAGGAGTGAAG	480		
QY	1329	CTGAGACCCACATACGTGTCCAGCATCAGAGCTGGCGGCTCATTTGGCAAGAGTGA	1388		
DB	481	CTGAGACCCACATACGTGTCCAGCATCAGAGCTGGCGGCTCATTTGGCAAGAGTGA	540		
QY	1389	AAACGGTGAACGAGTTCAGAAATTTGACGGCAGCTGAGGTGGTAGTACCAAGACCCAG	1448		
DB	541	AAACGGTGAACGAGTTCAGAAATTTGACGGCAGCTGAGGTGGTAGTACCAAGACCCAG	600		
QY	1449	ACCCCTGATGAAACGACAGCTCATCTGTAATAATCATCGGACATTTCTATGCCAGTCA	1508		
DB	601	ACCCCTGATGAAACGACAGCTCATCTGTAATAATCATCGGACATTTCTATGCCAGTCA	660		
QY	1509	ATGGCTCAACGGAAGATCCGACATCTTGGCCCGCAGTTAAGCAGCAGCATCAGAGGGA	1568		
DB	661	ATGGCTCAACGGAAGATCCGACATCTTGGCCCGCAGTTAAGCAGCAGCATCAGAGGGA	720		
QY	1569	CAGAGTAAACAGGCCAGCAGGAGGAGTGAACAGCCCTCCCTGTCTCCCTTNGAGTCC	1628		
DB	721	CAGAGTAAACAGGCCAGCAGGAGGAGTGAACAGCCCTCCCTGTCTCCCTTNGAGTCC	780		
QY	1629	AGGACAAACCGGCGCAATCGAGTGTGCTCTCCCGCAGGCTGAGATGATGG	1688		
DB	781	AGGACAAACCGGCGCAATCGAGTGTGCTCTCCCGCAGGCTGAGATGATGG	840		
QY	1689	GAATCCGGGACANTGGCGGGCTGTAGATGAGTTTGGCCACTTGATTGAGAAAGATG	1748		
DB	841	GAATCCGGGACANTGGCGGGCTGTAGATGAGTTTGGCCACTTGATTGAGAAAGATG	900		

QY	1749	TTCCAGTGAGGAACCTGATCTTNTCAGCCCCCAACCCCAACCCCAACCTGT	1808		
DB	901	TTCCAGTGAGGAACCTGATCTTNTCAGCCCCCAACCCCAACCCCAACCTGT	960		
QY	1809	NTGCCCTCGGGGTGTCAGAAATTTAGCGCAAGCAGCTTTTAAAGCTGGATTGTTAAA	1868		
DB	961	CTGCCCTCGGGGTGTCAGAAATTTAGCGCAAGCAGCTTTTAAAGCTGGATTGTTAAA	1020		
QY	1869	GAACTCTCAGGCCCCCAGAGAGGCTGATCACAACCTCAGTGGGAGAAAAATAAAT	1928		
DB	1021	GAACTCTCAGGCCCCCAGAGAGGCTGATCACAACCTCAGTGGGAGAAAAATAAAT	1080		
QY	1929	TTCTTTTCAAGTTTAAAA	1946		
DB	1081	TTCTTTTCAAGTTTAAAA	1098		
RESULT 8					
ABX73907					
ID	ABX73907	standard; DNA; 1129 BP.			
XX	ABX73907;				
XX	XX				
DT	18-MAR-2003	(first entry)			
XX	Human novel polynucleotide #735.				
DE	Human; gene; ds; neural disorder; immune system disorder; renal disorder;				
XX	muscular disorder; respiratory disease; reproductive disorder;				
KW	gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;				
KW	hyperproliferative disorder; inflammatory disease; allergic reaction;				
KW	blood related disorder; cancer; immunosuppressive; antiinflammatory;				
KW	cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;				
KW	haemostatic; antiarteriosclerotic.				
XX	Homo sapiens.				
OS	US2002132753-A1.				
PN	19-SEP-2002.				
PD	17-JAN-2001; 2001US-00764864.				
XX	31-JAN-2000; 2000US-0179065P.				
PR	04-FEB-2000; 2000US-0180628P.				
PR	28-JUN-2000; 2000US-0214886P.				
PR	07-JUL-2000; 2000US-0216647P.				
PR	11-JUL-2000; 2000US-0217487P.				
PR	11-JUL-2000; 2000US-0217496P.				
PR	14-JUL-2000; 2000US-0218290P.				
PR	26-JUL-2000; 2000US-0220963P.				
PR	14-AUG-2000; 2000US-0224518P.				
PR	14-AUG-2000; 2000US-0224519P.				
PR	14-AUG-2000; 2000US-0225267P.				
PR	14-AUG-2000; 2000US-0225268P.				
PR	14-AUG-2000; 2000US-0225270P.				
PR	14-AUG-2000; 2000US-0225447P.				
PR	14-AUG-2000; 2000US-0225757P.				
PR	14-AUG-2000; 2000US-0225758P.				
PR	22-AUG-2000; 2000US-0226868P.				
PR	30-AUG-2000; 2000US-022824P.				
PR	01-SEP-2000; 2000US-0228287P.				
PR	01-SEP-2000; 2000US-0229343P.				
PR	01-SEP-2000; 2000US-0229344P.				
PR	01-SEP-2000; 2000US-0229345P.				
PR	05-SEP-2000; 2000US-0229509P.				
PR	05-SEP-2000; 2000US-0229513P.				
PR	08-SEP-2000; 2000US-0231413P.				
PR	21-SEP-2000; 2000US-0234223P.				
PR	21-SEP-2000; 2000US-0234274P.				
PR	25-SEP-2000; 2000US-0234997P.				

Db 378 ACCACCGAAGACCTGACTCCTAAAGTTCTATGGTTATCATCTGAGCCGACAGGC 437
QY 1253 CMAATTCAGGCTCAGGGAAGAAATCTATGCAAACTCAAGGAGGAAGAACTTCTTTGGTCC 1312
Db 438 CCAATTCAGGCTCAGGGAAGAAATCTATGCAAACTCAAGGAGGAAGAACTTCTTTGGTCC 497
QY 1313 CAAGGAGGAAGTGAAGCTGAGACCCACATACCTGTGCGGAGCATCAGCAGCTGGCGGGT 1372
Db 498 CAAGGAGGAAGTGAAGCTGAGACCCACATACCTGTGCGGAGCATCAGCAGCTGGCGGGT 557
QY 1373 CATTCGCAAGGTGGAAAAACGGTGAACAGATTGCAAGATTCACGCGCAGCTCAGGTGCT 1432
Db 558 CATTCGCAAGGTGGAAAAACGGTGAACAGATTGCAAGATTCACGCGCAGCTCAGGTGCT 617
QY 1433 AGTACCAAGAGACAGACCCCTGATGAGAACGACAGGTCATCGTGAATAATCATCGGACA 1492
Db 618 AGTACCAAGAGACAGACCCCTGATGAGAACGACAGGTCATCGTGAATAATCATCGGACA 677
QY 1493 TTTCTATGCCAGTCAGATGCTCAACGGAAGATCCGAGACATCCTGGCCAGGTTAAGCA 1552
Db 678 TTTCTATGCCAGTCAGATGCTCAACGGAAGATCCGAGACATCCTGGCCAGGTTAAGCA 737
QY 1553 GCAGCATCAGAAGGACAGAGTAACAGGCCAGGACGAGGAAGTGACACGCGCTCC 1612
Db 738 GCAGCATCAGAAGGACAGAGTAACAGGCCAGGACGAGGAAGTGACACGCGCTCC 797
QY 1613 CTGTCCTTNGATCCAGGACAAACGAGGAGAAATCGAGAGTGCTCTCCCGGCGAG 1672
Db 798 CTGTCCTTNGATCCAGGACAAACGAGGAGAAATCGAGAGTGCTCTCCCGGCGAG 857
QY 1673 GCCTGAGATGATGGGAATCCGGACACNTGGCGGCTGTAGATCAGGTTTGGCCAC 1732
Db 858 GCCTGAGATGATGGGAATCCGGACACNTGGCGGCTGTAGATCAGGTTTGGCCAC 917
QY 1733 TTGATGTGAAGATGTTTCCAGTGAGGAACCTGATCTNTAGCCGCCAAACACCCACCA 1792
Db 918 TTGATGTGAAGATGTTTCCAGTGAGGAACCTGATCTNTAGCCGCCAAACACCCACCA 977
QY 1793 ATTGCCCCAACACTGNTGCCCTCGGGTGTCAGAAATNTAGCGAAGGCACTTTTAA 1852
Db 978 ATTGCCCCAACACTGNTGCCCTCGGGTGTCAGAAATNTAGCGAAGGCACTTTTAA 1037
QY 1853 ACGTGGAATGTTTAAAGAGCTCTCAGGCCCCACCAAGAGGCTGATCACAACCTCAGTG 1912
Db 1038 ACGTGGAATGTTTAAAGAGCTCTCAGGCCCCACCAAGAGGCTGATCACAACCTCAGTG 1946
QY 1913 GGAAGAAAAATAAATTTCTTCAGGTTTAAAA 1946
Db 1098 GGAAGAAAAATAAATTTCTTCAGGTTTAAAA 1131

RESULT 6

ABX73489

ID ABX73489 standard; DNA; 1182 BP.

AC AC

XX ABX73489;

DT 18-MAR-2003 (first entry)

XX 18-MAR-2003 (first entry)

DE Human novel polynucleotide #317.

XX Human; gene; ds; neural disorder; immune system disorder; renal disorder;

KW muscular disorder; respiratory disease; reproductive disorder;

KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;

KW hyperproliferative disorder; inflammatory disease; allergic reaction;

KW blood related disorder; cancer; immunosuppressive; antiinflammatory;

KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;

KW haemostatic; antiarteriosclerotic.

XX Homo sapiens.

OS Homo sapiens.

XX US2002132753-A1.

EN US2002132753-A1.

XX PD
XX PF
XX XX
XX PR 19-SEP-2002.
XX PR 17-JAN-2001; 2001US-00764864.
XX PR 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 28-JUN-2000; 2000US-0214886P.
XX PR 07-JUL-2000; 2000US-0216647P.
XX PR 07-JUL-2000; 2000US-0216880P.
XX PR 11-JUL-2000; 2000US-0217487P.
XX PR 11-JUL-2000; 2000US-0217496P.
XX PR 14-JUL-2000; 2000US-0218290P.
XX PR 26-JUL-2000; 2000US-0220963P.
XX PR 26-JUL-2000; 2000US-0220964P.
XX PR 14-AUG-2000; 2000US-0224518P.
XX PR 14-AUG-2000; 2000US-0224519P.
XX PR 14-AUG-2000; 2000US-0225267P.
XX PR 14-AUG-2000; 2000US-0225268P.
XX PR 14-AUG-2000; 2000US-0225270P.
XX PR 14-AUG-2000; 2000US-0225447P.
XX PR 14-AUG-2000; 2000US-0225757P.
XX PR 14-AUG-2000; 2000US-0225758P.
XX PR 22-AUG-2000; 2000US-0228688P.
XX PR 30-AUG-2000; 2000US-0228924P.
XX PR 01-SEP-2000; 2000US-0229287P.
XX PR 01-SEP-2000; 2000US-0229343P.
XX PR 01-SEP-2000; 2000US-0229344P.
XX PR 01-SEP-2000; 2000US-0229345P.
XX PR 05-SEP-2000; 2000US-0229509P.
XX PR 05-SEP-2000; 2000US-0229513P.
XX PR 08-SEP-2000; 2000US-0231413P.
XX PR 21-SEP-2000; 2000US-0234223P.
XX PR 21-SEP-2000; 2000US-0234274P.
XX PR 25-SEP-2000; 2000US-0234997P.
XX PR 27-SEP-2000; 2000US-0235834P.
XX PR 29-SEP-2000; 2000US-0236327P.
XX PR 29-SEP-2000; 2000US-0236367P.
XX PR 29-SEP-2000; 2000US-0236368P.
XX PR 29-SEP-2000; 2000US-0236369P.
XX PR 29-SEP-2000; 2000US-0236370P.
XX PR 02-OCT-2000; 2000US-0236802P.
XX PR 02-OCT-2000; 2000US-0237037P.
XX PR 02-OCT-2000; 2000US-0237038P.
XX PR 02-OCT-2000; 2000US-0237039P.
XX PR 02-OCT-2000; 2000US-0237040P.
XX PR 13-OCT-2000; 2000US-0239935P.
XX PR 20-OCT-2000; 2000US-0240960P.
XX PR 20-OCT-2000; 2000US-0241785P.
XX PR 20-OCT-2000; 2000US-0241809P.
XX PR 01-NOV-2000; 2000US-0244617P.
XX PR 17-NOV-2000; 2000US-0249299P.
XX PR 08-DEC-2000; 2000US-0251856P.
XX PR 08-DEC-2000; 2000US-0251868P.
XX PR 08-DEC-2000; 2000US-0251869P.

(ROSE/) ROSEN C A.

(RUBE/) RUBEN S M.

(BARA/) BARASH S C.

Rosen CA, Ruben SM, Barash SC;

WPI; 2003-147444/14.

P-PSDB; ABUS5229.

New polypeptides and nucleic acids, useful in gene therapy for treating,

inhibiting or preventing e.g. neural, immune system, muscular,

respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or

renal disorders.

Claim 1; SEQ ID NO 327; 402pp; English.

The invention relates to human novel polypeptides and their associated

CC

Db 1385 GCATCATTCGCAAGAGGGCCACACATCAACACATCTCCCTTTCCGCGAGCGCTCC 1444
QY 1182 ATCAAGATTGACACACCGGAAACACCTGACTCCAAAGTTGGTATGATCATCATCTGGA 1241
Db 1445 ATCAAGATTGCTCCACGAGAAACACCTGACTCCAAAGTTGGAATGGTCATCATCTGGA 1504
QY 1242 CCGCCAGAGGCCCAATTCAGAGCTCAGGAGAGATCTATGGCAAACTCAAGGAGGAAC 1301
Db 1505 CCCCCAGAGGCTCAGTTCAAGGCTCAGGAGAAATTTATGGCAAACTTAAAGAGGAAT 1564
QY 1302 TTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACACATACGTGTGCCAGCATCAGCA 1361
Db 1565 TTCTTTGGTCCCAAGGAGGAAGTGAAGCTAGAGACCCACACATACGGGTTCGGCTTCAGCA 1624
QY 1362 GCTGGCGGGTCTATGGCAAGGTGGAAACCGGTGAACCGATTGACAGAAATTTGACGGCA 1421
Db 1625 GCCGGCGGCTATCGCGCAAGGCGCAAAACCGTGAATGAGCTGCAGAACTTTGACTGCA 1684
QY 1422 GCTGAGGTGCTAGTACCAAGAGACCAAGACCCCTGATGAGAACCAAGCTCATCTGTA 1481
Db 1685 GCTGAGGTGCTAGTACCAAGAGACCAAGACCCCGATGAGAACCAAGCTCATCTGTA 1744
QY 1482 ATCATCGGACATTTATGCCATGCTAGATGGCTCAACGGGAAGATCCGAGACATCTGGCC 1541
Db 1745 ATCATCGGACATTTATGCCATGCTAGATGGCTCAGCGGAGATCCGAGACATCTGGCT 1804
QY 1542 CAGGTTAAGCAGCAGCATCAGAGGACAGAGTAACAGGCCAGGACCGAGGAAGTCA 1601
Db 1805 CAAGTTAAGCAACAGCAGGAGGACAGCAACCTGGCCAGGACCGGAGGAAGTCA 1864
QY 1602 -CCAGCCCTCCCTGTCCTTGAGTCCAGGACCAACAGGGCAGAA----- 1647
Db 1865 CCCC GCCCTCTCTGCTCCATTTGGCTCCAGATCAGCAGGAGAACACAGAACTCGAGGG 1924
QY 1648 -----ATCAGAGTGTGCTCTCCCGGAGGCTGAGATGAGTGGGAATCCGGGAC 1699
Db 1925 GCGGGTGGAGGCGCGGTGTGTTTTCCAGCAGGCTGAGATGAGTGGGAATCAGGG-C 1983
QY 1700 ACNTGGCGGGCTGTAGATCAGGTTGGCCACTTGTGAGAAAGATTTCCAGTGAGG 1759
Db 1984 ATTTGGGCTGCTGAGATCAGGTTGACACTGTATTGAGAACAAATGTTCCAGTGAGG 2043
QY 1760 AACCTGTATCTNTACGCCCAACCAACCACTGGCCCAACACTGNTGCCCTCGG 1819
Db 2044 AATCTGTATCTCTGCCCCCAA--TTGAGCCAGCTGGCCACAGCCACCCCTTGGAAAT 2101
QY 1820 GGTGTCAAAATNTAGCGCAAGGCATTTTAAACGTGGATTTTAAAGAGCTCTCCA 1879
Db 2102 CACCAATTCATCATAGCTTGGTTGCTTTTAAACGTGGATTTGCT--TGAAGTTCTCCA 2159
QY 1880 GGCCCAACCAAGAGGCTGATCAGACCTCAGTGGGAAGAAATTAATTTCTTCAGGT 1939
Db 2160 GCCTCCATGGAAGATGGTCCAGTCCAGTGGGGAAGAGAAATTAATTTCTTCAGGT 2219
QY 1940 TTTA 1943
Db 2220 TTTA 2223

RESULT 5

AAS26148
ID AAS26148 standard; cDNA; 1182 BP.
XX
AC AAS26148;
XX
XX
DT 07-NOV-2001 (first entry)
XX
DE Human cDNA encoding a novel secreted protein, Seq ID 327.
XX
KW Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;

secreted protein; rheumatoid arthritis; hyperproliferative disorder;
cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
cerebral ischaemia; angiogenesis; nervous system disorder;
Alzheimer's disease; infection; ocular disorder; corneal infection;
wound healing; epithelial cell proliferation; skin ageing; food additive;
preservative; antiproliferative.
Homo sapiens.
WO200155322-A2.

XX 02-AUG-2001.
XX
PN 17-JAN-2001; 2001WO-US0001341.
XX
PD 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216800P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.

Mon Jul 26 09:17:14 2004

Db 1081 CCCTGTCCTTTCAGAGTCCAGGACAAACAGGCGAGAAATCGAGAGTGTGCTCTCCCGGC 1140
 Qy 1671 AGGCTGAGATGAGTGGGATCCGGGACACNTGGCGGCTGTAGATCAGGTTTGCC 1730
 Db 1141 AGGCTGAGATGAGTGGGATCCGGGACACCTGGCGGCTGTAGATCAGGTTTGCC 1200
 Qy 1731 ACTTCATTGAGAAAGATGTTCCAGTGGAGAACCTTGATCTNTCAGCCCCAAACACCCACC 1790
 Db 1201 ACTTCATTGAGAAAGATGTTCCAGTGGAGAACCTTGATCTNTCAGCCCCAAACACCCACC 1260
 Qy 1791 CAATTGGCCCAACACCTGTTGCGGCTCGGGGTGTGAGAAATTTAGCGCAAGCACTTTT 1850
 Db 1261 CAATTGGCCCAACACCTGTTGCGGCTCGGGGTGTGAGAAATTTAGCGCAAGCACTTTT 1320
 Qy 1851 AAACGTGGATTTTAAAGAGCTCTCCAGGCCCCACCAAGAGGTTGATCACACCTCAG 1910
 Db 1321 AAACGTGGATTTTAAAGAGCTCTCCAGGCCCCACCAAGAGGTTGATCACACCTCAG 1380
 Qy 1911 TGGGAGAGAAATAAATAATTCCTTCAGGTTTAAAA 1946
 Db 1381 TGGGAGAGAAATAAATAATTCCTTCAGGTTTAAAA 1416

RESULT 4
 AAZ10617
 ID AAZ10617 standard; cDNA; 2224 BP.
 XX AAZ10617;
 XX AC
 XX DT 17-NOV-1999 (first entry)
 XX cDNA encoding a murine c-myc coding region determinant binding protein.
 DE c-myc coding region determinant binding protein; CRD-BP; tumor; c-myc;
 XX endonucleolytic attack; half-life; breast cancer; colon cancer;
 KW pancreatic cancer; ss.
 KW Mus musculus.

XX Key Location/Qualifiers
 XX 131..1864
 XX CDS /*tag= a
 XX
 XX W09946594-A2.
 XX 16-SEP-1999.
 XX 05-MAR-1999; 99WO-US004897.
 XX 09-MAR-1998; 98US-0077372P.
 XX (WISC) WISCONSIN ALUMNI RES FOUND.
 XX Ross J;
 XX WPI; 1999-551506/46.
 XX P-F5DB; AAY30849.
 XX Diagnosing presence or absence of a tumor in a human by examining c-myc
 XX coding region determinant-binding protein.
 XX Example; Fig 1A-D; 79pp; English.
 XX The present sequence encodes a murine c-myc coding region determinant
 XX binding protein (CRD-BP). The presence or absence of a tumor can be
 XX determined by determining the levels of CRD-BP present in the suspect
 XX tissue, where the CRD-BP shields c-myc RNA from endonucleolytic attack
 XX and so prolongs its half-life. The methods are used for diagnosing
 XX presence or absence of a tumor in a human, especially breast, colon and
 XX pancreatic cancer. They are also used to inhibit cancer cell growth
 XX Sequence 2224 BP; 580 A; 609 C; 618 G; 417 T; 0 U; 0 Other;
 XX SQ

Query Match 66.7%; Score 1297.8; DB 2; Length 2224;
 Best Local Similarity 83.4%; Pred. No. 0;
 Matches 1554; Conservative 0; Mismatches 277; Indels 33; Gaps 6;
 Qy 108 AGGAGCCGGAATTCATTAATCCGAAATATTCACCCAGCTCCGATGGAAGTACTGGAC 167
 Db 365 AGGAGTCCGAAATACAGATCCGCAATATTCACCTCAGCTCCGATGGAAGTCTAGAT 424
 Qy 168 AGCCTGCTGGCTCAGTATGTTAGTACAGTGTGAGAACTGTGAGCAAGTGAACACCTGAG 227
 Db 425 AGCCTGCTGGCTCAGTATGTTAGTACAGTGTGAGAACTGTGAGCAAGTGAACACCTGAG 484
 Qy 228 AGGAGTGTGCTGCAATGTCACTTATTCACACCGGGAGCAGACAGGCAAGCATCATCAAG 287
 Db 485 ACAGCGTGTGTCTCACTGCTCACTTCTAACCAGGAGCAGACAGGCAAGCATCATCAAG 544
 Qy 288 CTGAATGGCCACACAGTTGGAGAACATGCTTCAAGTCTTCTACATCCCGATGAGCAG 347
 Db 545 CTAAATGGCCATCAACTGGAGAACCATGCTTCAAGTCTTCTACATACCTGATGAGCAG 604
 Qy 348 ATAGCACAGGACCTCAGATGGCGCGCGAGGGGCTTTGGCTCTCGGGGTACGCCCCG 407
 Db 605 ATAAACCAAGGTCCTGAGAAATGGCGCTGCTGAGAGGCTTTGGGTCTCGGGCCAGCCCCG 664
 Qy 408 CAGGCTCACCTGTGCGAGCGGGGCCCCAGCCAGCAGCAGCAGTGTGACATCCCCCTT 467
 Db 665 CAGGGTCCGCTGCGAGCAGGGGCTCCAGCCAAAGCAGCAGCAGTGTGACATCCCCCTT 724
 Qy 468 CGGCTCCTGTGCTCCACCCAGTATGTTGGTGCATTTATTCGCAAGGAGGGGCCACCATC 527
 Db 725 CGGCTCCTGTGCTCCACCCAGTATGTTGGTGCATTTATTCGCAAGGAGGGGCCACCATC 784
 Qy 528 CGCAACATCAAAACAGACCCAGTCCAGATAGACGTGATGATGAGAGGAGAAACGAGGT 587
 Db 785 CGAAACATCAAAACAGACCCAGTCCAGATAGACGTGATGATGAGAGGAGAAACGAGGT 844
 Qy 588 GCAGCTGAAAAGCCATCAGTGTGCTCCACCCCTGAGGGGTGCTCTCGCTTGTAAAG 647
 Db 845 GCTCGGAGAGGCCATCAGCTGCATTCACCCCTGAAAGCTGTCTCTCGCGGTGCAAG 904
 Qy 648 ATGATCTTGGAGATTATGATTAAGAGGCTTAAGACACCAACACCGCTGACGAGTTTCCC 707
 Db 905 ATGATCTTGGAGATTATGATTAAGAGGCTTAAGACACCAACACCGCTGACGAGTTTCCC 964
 Qy 708 CTGAGATCTTGGCCCAATAATACTTTGTAGGGGTCTCTCATTTGGCAAGAGGAGCGAAC 767
 Db 965 CTGAAGATCTTGGCTCATPAACATCTTCTCGGGGACTCATTTGGCAAGAGGCGCGAAC 1024
 Qy 768 CTGAAGAGGTAGACCAAGATACCGAGACAAATTCACATCTCTCTGTTGCAAGACCTT 827
 Db 1025 CTGAAGAGGTAGACCAAGATACCGAGACAAATTCACATCTCTCTGTTGCAAGACCTT 1084
 Qy 828 ACCCTTTACACCTTGAGAGGACCATCACTGTGAAGGGGGCCCATCGAATTTGTTGAGG 887
 Db 1085 AGGCTCTATAACCTTGAGAGGACCATCACTGTGAAGGGGGCCCATTTGAGAACTGTTG 1144
 Qy 888 GCGGAGCGGAAATTAATGAAGAAAGTTGCGGAGGCTTATGAGAACTGTTGAGGCTG 947
 Db 1145 GCGGAGCGGAGATCATGAAGAAAGTTGAGAGGCTTACGAGAACGACGCTGGCGCGCATG 1204
 Qy 948 AGC-----TCTCAGCTGATCCCTGGCCCTGAAACCTGGCTGTGTAGGTCTTTTCCAGCT 1001
 Db 1205 AGCTTGCACTCCCACTCATCTCTGGCTTAACTGGCTGTGTAGGTCTTTTCCAGCT 1264
 Qy 1002 TCATCAGCGAGTCCCGCGGCTCCAGCAGGCTTACTGGGCTCTCTCCCTATAGCTCC 1061
 Db 1265 TCATCAGCGGCTGCTCCCTCTCTCCAGCAGGCTGCTGAGGCTGCTCTCCCTATAGCTCC 1324
 Qy 1062 TTTATCAGGCTCCCGCAGCAGGAGATGTGAGGCTTTTATCCCCCAGGAGGAGTGGGC 1121
 Db 1325 TTCATCAGGCTCCCGCAGCAGGAGATGTGAGGCTTTTATCCCCCAGGAGGAGTGGGC 1384
 Qy 1122 GCCATCATCGGCAAGAGGCGGAGCAGCATCAACAGCTCTCCCGGTTTGCCAGGCGCTCC 1181


```
Db 1470 CGGCTGTAGATCAGGTTTGCCTACTGANTGAGAAAGATGTTCCAGTGGAGACCTGGA 1529
Qy 1768 TCINTCAGCCCCCAACACCCACCAATTTGGCCCAACACTGNTGCCCTCGGGGTGTCAG 1827
Db 1530 TCINTCAGCCCCCAACACCCACCAATTTGGCCCAACACTGNTGCCCTCGGGGTGTCAG 1589
Qy 1828 AAATNTAGCGCAGGACACTTTTAAACGTGGATTTTAAAGAGAGCTCTCCAGGCCCCAC 1887
Db 1590 AAATNTAGCGCAGGACACTTTTAAACGTGGATTTTAAAGAGAGCTCTCCAGGCCCCAC 1649
Qy 1888 CAAGAGGTCGATCAGACACTTCACTTGGGAGAGAAATAAATTTCCCTTTCAGGTTTAAAA 1946
Db 1650 CAAGAGGTCGATCAGACACTTCACTTGGGAGAGAAATAAATTTCCCTTTCAGGTTTAAAA 1708

RESULT 3
AAK94782
ID AAK94782 standard; cDNA; 2780 BP.
XX
AC AAK94782;
XX
XX
DT 06-NOV-2001 (first entry)
DE Human full-length cDNA, SEQ ID NO: 3886.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
OS Homo sapiens.
XX
XX EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-00114089.
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX
XX
PA (HELI-) HELIX RES INST.
XX
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
XX WPI; 2001-524255/58.
DR P-PSDB; AAM93826.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their use
in genetic manipulation.
XX
XX Claim 8; SEQ ID NO 3886; 1380pp + Sequence Listing; English.
XX
XX The invention relates to primers for synthesising full length cDNA
XX clones. 830 cDNA molecules encoding a human protein have been isolated
XX and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
XX been determined. Primers for synthesising the full length cDNA are useful
XX for clarifying the function of the protein encoded by the cDNA. The full
XX length clones were obtained by construction of full length enriched cDNA
XX libraries that were synthesised by the oligo-capping method. The primers
XX enable the production of the full length cDNA easily without any special
XX methods. The present sequence is a full length human cDNA of the
XX invention. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in CD-ROM format directly
XX from EPO
XX
XX Sequence 2780 BP; 768 A; 681 C; 671 G; 660 T; 0 U; 0 Other;
XX
XX
XX Query Match 71.4%; Score 1389; DB 4; Length 2780;
XX Best Local Similarity 99.2%; Pred. No. 0;
XX Matches 1405; Conservative 0; Mismatches 5; Indels 6; Gaps 1;
XX
XX 537 ACAAAACAGACCCCAAGATAGAGTGCATAGGAGAGAGAACGAGGTGCAGCTGAA 596
```

```
Db 1 ACAAACAGACCCCAAGATAGAGTGCATAGGAGAGAGAACGAGGTGCAGCTGAA 60
Qy 597 AAAGCCATCAGTGTGCACTCCACCCCTGAGGCTGCTCCTCCGCTGTGAAGATGATCTTG 656
Db 61 AAAGCCATCAGTGTGCACTCCACCCCTGAGGCTGCTCCTCCGCTGTGAAGATGATCTTG 120
Qy 657 GAGATTATGATGAAGAGGCTAAGGACACCAAAACGGCTGACGAGGTTCCCTCGAAGATC 716
Db 121 GAGATTATGATGAAGAGGCTAAGGACACCAAAACGGCTGACGAGGTTCCCTCGAAGATC 180
Qy 717 CTGGCCCAATAAATCTTTGTAGGGCTCTCATNTGCAAGAGAGACCGAACCCTGAAGAAG 776
Db 181 CTGGCCCAATAAATCTTTGTAGGGCTCTCATNTGCAAGAGAGACCGAACCCTGAAGAAG 240
Qy 777 GTAGAGCAAGATACCGAGACAAATAACCATCTCCTCGTTGCAAGACCTTACCTTTTAC 836
Db 241 GTAGAGCAAGATACCGAGACAAATAACCATCTCCTCGTTGCAAGACCTTACCTTTTAC 300
Qy 837 AACCTTGAGAGACCATCACTGTGAAGGGGGCCATCGAGAAATTTGTGCAAGGCCGAGCAG 896
Db 301 AACCTTGAGAGACCATCACTGTGAAGGGGGCCATCGAGAAATTTGTGCAAGGCCGAGCAG 360
Qy 897 GAAATAATGAAGAAAGTTCCGGAGGCCCTATGAGAAATGATGTGGCTGCCATGAGCCTGAC 950
Db 361 GAAATAATGAAGAAAGTTCCGGAGGCCCTATGAGAAATGATGTGGCTGCCATGAGCCTGAC 420
Qy 951 TCTCACTGATCCCTGGCTGAAACCTGCTGCTGAGGCTGCTCCTATAGCTCCTTTATGAG 1070
Db 421 TCTCACTGATCCCTGGCTGAAACCTGCTGCTGAGGCTGCTCCTATAGCTCCTTTATGAG 540
Qy 1011 GCAGTCCCGCGCTCCAGCAGCGTTACTGGGGCTGCTCCTATAGCTCCTTTATGAG 1070
Db 481 GCAGTCCCGCGCTCCAGCAGCGTTACTGGGGCTGCTCCTATAGCTCCTTTATGAG 540
Qy 1071 GCTCCCGAGCAGGATGCTGAGTGTGTTATCCCGCCGAGGCTGAGTGGGCCCATCATC 1130
Db 541 GCTCCCGAGCAGGATGCTGAGTGTGTTATCCCGCCGAGGCTGAGTGGGCCCATCATC 600
Qy 1131 GCGAAGAGGGGCGAGCAGATCAAAACAGCTCTCCCGTTTCCAGGCTCCATCAGATT 1190
Db 601 GCGAAGAGGGGCGAGCAGATCAAAACAGCTCTCCCGTTTCCAGGCTCCATCAGATT 660
Qy 1191 GCACACCCGAAACACCTGACTCCAAAGTTGATGTTATCATCACTGAGACCGCCAG 1250
Db 661 GCACACCCGAAACACCTGACTCCAAAGTTGATGTTATCATCACTGAGACCGCCAG 720
Qy 1251 GCCCAATTCAGGCTCAGGAGAGATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGT 1310
Db 721 GCCCAATTCAGGCTCAGGAGAGATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGT 780
Qy 1311 CCCAAGAGGAGAGTGAAGCTGGAGACCCACATAGCTGTGCCAGCATCAGAGCTGGCGG 1370
Db 781 CCCAAGAGGAGAGTGAAGCTGGAGACCCACATAGCTGTGCCAGCATCAGAGCTGGCGG 840
Qy 1371 GTCATTTGGCAAAAGTGGAAACCGTGAACCGAGTTGAGAAATTTGACGGCAGCTGAGTG 1430
Db 841 GTCATTTGGCAAAAGTGGAAACCGTGAACCGAGTTGAGAAATTTGACGGCAGCTGAGTG 900
Qy 1431 GTAGTACCAAGAGACAGACCCCTGATGAGAACGACAGCTCATCGTGAATAATCATCGGA 1490
Db 901 GTAGTACCAAGAGACAGACCCCTGATGAGAACGACAGCTCATCGTGAATAATCATCGGA 960
Qy 1491 CATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCTGCGCCAGGTTAG 1550
Db 961 CATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCTGCGCCAGGTTAG 1020
Qy 1551 CAGCAGCATCAGAGGGAAGAGTAAACAGGCCCGAGGCGAGGAGTGAACAGAGCCCT 1610
Db 1021 CAGCAGCATCAGAGGGAAGAGTAAACAGGCCCGAGGCGAGGAGTGAACAGAGCCCT 1080
Qy 1611 CCTGTCCCTTNGTCCAGGACAAACAGGGCAGAAATCGAGAGTGTCTCTCCCGGC 1670
```

XX	11-FEB-2000	(first entry)	
XX	DNA encoding cancer associated antigen KOC-2.		
XX	Cancer associated antigen; KOC-2; cancer; vaccine; CT7; ss.		
XX	Homo sapiens.		
XX	WO9954738-A1.		
XX	28-OCT-1999.		
XX	16-MAR-1999;	99WO-US0005766.	
XX	17-APR-1998;	98US-00061709.	
XX	(LUDW-) LUDWIG INST CANCER RES.		
XX	Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old LJ;		
XX	WPI; 2000-013284/01.		
XX	Nucleotides representing cancer-associated genes, used to develop		
XX	products for the diagnosis, monitoring and treatment of cancers.		
XX	Claim 55; Page 40; 44pp; English.		
XX	The present sequence represents a cancer associated antigen gene		
XX	designated KOC-2. The specification also describes a cancer associated		
XX	antigen designated CT7. The CT7 polynucleotide was isolated from SK-MEL-		
XX	37 melanoma cells. The polypeptide has some homology with MAGS-10,		
XX	the protein has a repetitive pattern, with repeats rich in serine,		
XX	proline, glutamine and leucine, and an almost invariable core of the		
XX	peptide given in MAY3877. The CT7 polypeptide can be processed to		
XX	peptides which provoke lysis by cytolytic T cells. The polynucleotides		
XX	and polypeptides can be used for treating a cancerous condition and		
XX	screening for or diagnosing cancerous conditions. The cancer associated		
XX	antigens can be used as an immunogenic or vaccine composition with an		
XX	adjuvant, e.g. a cytokine, a saponin, or granulocyte macrophage-colony		
XX	stimulating factor (GM-CSF)		
SQ	Sequence 1708 BP; 447 A; 469 C; 473 G; 314 T; 0 U; 5 Other;		
	Query Match	85.8%;	Score 1669.2; DB 3; Length 1708;
	Best Local Similarity	99.8%;	Pred. No. 0;
	Matches 1676;	Conservative 0;	Mismatches 3; Indels 0; Gaps 0;
Qy	268	CCAGGCAAGCCATCATCAAGCTGAATGGCCACCACTGGAGAACCATGCCCTGAAGGTCT	327
Db	30	CCCGGGGAGCCATCATGAAGCTGAATGGCCACCACTGGAGAACCATGCCCTGAAGGTCT	89
Qy	328	CCTACATCCCGATGAGCAGATAGCAGGACCTGGAATGGGGCCGAGGGGCTTTG	387
Db	90	CCTACATCCCGATGAGCAGATAGCAGGACCTGGAATGGGGCCGAGGGGCTTTG	149
Qy	388	GCTCTCGGGGTACGCCCGCAGGGCTACCTGTGGCAGCGGGGCCCCAGCCAGCAGC	447
Db	150	GCTCTCGGGGTACGCCCGCAGGGCTACCTGTGGCAGCGGGGCCCCAGCCAGCAGC	209
Qy	448	AGCAAGTGGACATCCCGCTCGGCTCCGTGGTCCCGCCAGTATGTTGGGTGCCATTATG	507
Db	210	AGCAAGTGGACATCCCGCTCGGCTCCGTGGTCCCGCCAGTATGTTGGGTGCCATTATG	269
Qy	508	GCAAGAGGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCAGATAGACGTGC	567
Db	270	GCAAGAGGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCAGATAGACGTGC	329
Qy	568	ATAGAGAGAGAACCCAGGTGAGCTGAAAGAGCCATCAGTGTGCACTCCACCCCTGAGG	627
Db	330	ATAGAGAGAGAACCCAGGTGAGCTGAAAGAGCCATCAGTGTGCACTCCACCCCTGAGG	389

Qy	628	GCTGCTCTCCGCTTGTAAAGATGATCTTTGGAGATTATGCAATAAGAGGCTAAGGACCA	687
Db	390	GCTGCTCTCCGCTTGTAAAGATGATCTTTGGAGATTATGCAATAAGAGGCTAAGGACCA	449
Qy	688	AAACGGGTGACGAGGTTCCCTCTGAAGATCTCGGCCATAATAAATTTGTAGGGCGCTCA	747
Db	450	AAACGGGTGACGAGGTTCCCTCTGAAGATCTCGGCCATAATAAATTTGTAGGGCGCTCA	509
Qy	748	TTGGCAAGGAGGACGGAACTTGAAGAAGGTAGAGCAAGATACCGAGACAAAATACCA	807
Db	510	TTGGCAAGGAGGACGGAACTTGAAGAAGGTAGAGCAAGATACCGAGACAAAATACCA	569
Qy	808	TCTCTCTGTTGCAAGACCTTACCTTTTACAACCTTGAGAGGACCATCTACTGTGAAGGGG	867
Db	570	TCTCTCTGTTGCAAGACCTTACCTTTTACAACCTTGAGAGGACCATCTACTGTGAAGGGG	629
Qy	868	CCATCGAATTTGTTGCGAGGCGGAGAGGAAATATGAAGAAGTTGCGGGAGGCGCTATG	927
Db	630	CCATCGAATTTGTTGCGAGGCGGAGAGGAAATATGAAGAAGTTGCGGGAGGCGCTATG	689
Qy	928	AGAATGATGTGCTGCCATGAGCTCTCACCTCATCCCTGGCCTGAACCTTGGCTGTGTAG	987
Db	690	AGAATGATGTGCTGCCATGAGCTCTCACCTCATCCCTGGCCTGAACCTTGGCTGTGTAG	749
Qy	988	GTCTTTTCCAGCTTCATCCAGCGAGTCCCGCGCTCCCGAGCAGCTTACTGGGGCTG	1047
Db	750	GTCTTTTCCAGCTTCATCCAGCGAGTCCCGCGCTCCCGAGCAGCTTACTGGGGCTG	809
Qy	1048	CTCCCTATAGTCTCTTTATGCGAGGCTCCGAGCAGGAGATGTCGAGGTGTTTATCCCGG	1107
Db	810	CTCCCTATAGTCTCTTTATGCGAGGCTCCGAGCAGGAGATGTCGAGGTGTTTATCCCGG	869
Qy	1108	CCGAGCAGTGGCGGCATCATCGCAAGAGAGGGGAGCAGACATCAACAGCTCTCCCGT	1167
Db	870	CCGAGCAGTGGCGGCATCATCGCAAGAGAGGGGAGCAGACATCAACAGCTCTCCCGT	929
Qy	1168	TTCCGAGCGCTCCATCAGATTGCACACCCGAAACACTGACTCTCAAAGTTGCTATGG	1227
Db	930	TTCCGAGCGCTCCATCAGATTGCACACCCGAAACACTGACTCTCAAAGTTGCTATGG	989
Qy	1228	TTATCATCATGACCGCCAGAGCCCAATTAAGGCTCAGGGAGAGATCTATGCGCAAC	1287
Db	990	TTATCATCATGACCGCCAGAGCCCAATTAAGGCTCAGGGAGAGATCTATGCGCAAC	1049
Qy	1288	TCAAGGAGGAGAACTCTTTGTTCCCAAGGAGGAAGTGAAGCTGGAGCCACATACGTG	1347
Db	1050	TCAAGGAGGAGAACTCTTTGTTCCCAAGGAGGAAGTGAAGCTGGAGCCACATACGTG	1109
Qy	1348	TGCCAGCATCAGCAGCTGGCGGGTCAATTTGGCAAGGTGGAAGAACCGTGAACGAGTTGC	1407
Db	1110	TGCCAGCATCAGCAGCTGGCGGGTCAATTTGGCAAGGTGGAAGAACCGTGAACGAGTTGC	1169
Qy	1408	AGAAATTTGACGGCAGCTGAGTGTAGTACCAAGAGCCAGACCCCTCATGAGAACGACC	1467
Db	1170	AGAAATTTGACGGCAGCTGAGTGTAGTACCAAGAGCCAGACCCCTCATGAGAACGACC	1229
Qy	1468	AGGTTCATGCTGAATCATCGACATTTCTATGCGAGTCAGATGGCTCAACGGAAGATCC	1527
Db	1230	AGGTTCATGCTGAATCATCGACATTTCTATGCGAGTCAGATGGCTCAACGGAAGATCC	1289
Qy	1528	GAGACATCCTCGCCCGAGGTTAAGCAGCAGCATAGAGGGGACAGATACCCAGGCCGAGG	1587
Db	1290	GAGACATCCTCGCCCGAGGTTAAGCAGCAGCATAGAGGGGACAGATACCCAGGCCGAGG	1349
Qy	1588	CACGAGGAAGTGAACAGCCCTCCCTGCTCCCTTNGAGTCCAGGACCAACACCGGCGAGAA	1647
Db	1350	CACGAGGAAGTGAACAGCCCTCCCTGCTCCCTTNGAGTCCAGGACCAACACCGGCGAGAA	1409
Qy	1648	ATCGAGAGTGTCTCTCCCGGAGGCTGGAATGAGTGGGAATCCGGGACACNTGGGC	1707
Db	1410	ATCGAGAGTGTCTCTCTCCCGGAGGCTGGAATGAGTGGGAATCCGGGACACNTGGGC	1469
Qy	1708	CGGGCTGTAGATCAGGTTTGGCCACTTGATTGAGAAAGATGTTCCAGTGAGGAACCTGTA	1767

CC	macrophage-colony stimulating factor (GM-CSF)	
XX		
SQ	Sequence 1946 BP; 502 A; 528 C; 553 G; 358 T; 0 U; 5 Other;	
	Query Match 99.7%; Score 1941; DB 3; Length 1946;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 1946; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 GCTGTAGCGAGGGCTGGGGGCTGCTCTGTCTCCCTTCTTGGCGGCTGCGGCTCAGC	60
DB	1 GCTGTAGCGAGGGGCTGGGGGCTGCTCTGTCTCCCTTCTTGGCGGCTGCGGCTCAGC	60
QY	61 CCACCCAGAGGCGGGGTGGAGGGCGAGTGTCTAGCTTCCCGGGTTAGGAGCCGGA	120
DB	61 CCACCCAGAGGCGGGGTGGAGGGCGAGTGTCTAGCTTCCCGGGTTAGGAGCCGGA	120
QY	121 TTCAAAATCCGAATATTTCCACCCAGCTCCGATGGGAAGTACTGGACAGCTGTCTGGCTC	180
DB	121 TTCAAAATCCGAATATTTCCACCCAGCTCCGATGGGAAGTACTGGACAGCTGTCTGGCTC	180
QY	181 AGTATGTTACAGTAGAGAACTGTGAGCAAGTGAACACCCGAGAGTGAAGCGGCAAGTGTGA	240
DB	181 AGTATGTTACAGTAGAGAACTGTGAGCAAGTGAACACCCGAGAGTGAAGCGGCAAGTGTGA	240
QY	241 ATGTACCTTATTTCCAAACCGGGAGCAGACGAGCAAGCCATCTGAAGCTGAATGGCCACC	300
DB	241 ATGTACCTTATTTCCAAACCGGGAGCAGACGAGCAAGCCATCTGAAGCTGAATGGCCACC	300
QY	301 AGTTGAGAACCATGCTTCAAGGCTCTCTACATCCCGATGAGCAGATAGCAGAGGAC	360
DB	301 AGTTGAGAACCATGCTTCAAGGCTCTCTACATCCCGATGAGCAGATAGCAGAGGAC	360
QY	361 CTGAGATGGCGCGGAGGGGGCTTTGGCTCTCGGGTTCAGCCCGCAGGGGTCACTTG	420
DB	361 CTGAGATGGCGCGGAGGGGGCTTTGGCTCTCGGGTTCAGCCCGCAGGGGTCACTTG	420
QY	421 TGGCAGCGGGGCCCCAGCAGCAGCAGCAGTGGACATCCCCCTTCGCTCTCGCTGTC	480
DB	421 TGGCAGCGGGGCCCCAGCAGCAGCAGCAGTGGACATCCCCCTTCGCTCTCGCTGTC	480
QY	481 CCACCCAGTATGGGTGGGCTTATTTGCAAGGAGGGGCCACCATCGCAACATCAAA	540
DB	481 CCACCCAGTATGGGTGGGCTTATTTGCAAGGAGGGGCCACCATCGCAACATCAAA	540
QY	541 AACAGACCCAGTCCAAAGTAGACGTGCATAGGAAGGAGAACGCGAGTGCAGCTGAA	600
DB	541 AACAGACCCAGTCCAAAGTAGACGTGCATAGGAAGGAGAACGCGAGTGCAGCTGAA	600
QY	601 CCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCTCCGCTTGAAGATGATCTTGGAGA	660
DB	601 CCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCTCCGCTTGAAGATGATCTTGGAGA	660
QY	661 TTATGCAATAGAGGCTAAAGACACCAAAACGGCTGACGAGTTCCTTGAAGATCCTGG	720
DB	661 TTATGCAATAGAGGCTAAAGACACCAAAACGGCTGACGAGTTCCTTGAAGATCCTGG	720
QY	721 CCCATAATACTTTGTAGGGCTCTCATTTGGCAAGGAAGGACGGAACCTGAAGAGGTAG	780
DB	721 CCCATAATACTTTGTAGGGCTCTCATTTGGCAAGGAAGGACGGAACCTGAAGAGGTAG	780
QY	781 AGCAAGATACCGAGACAAATATCCATCTCTCGTTCAGACCTTACCTTTACAACC	840
DB	781 AGCAAGATACCGAGACAAATATCCATCTCTCGTTCAGACCTTACCTTTACAACC	840
QY	841 CTGAGGACCATCATCTGTGAAGGGGCCATCGAATTTGTGAGGGCCGAGCAGGAAA	900
DB	841 CTGAGGACCATCATCTGTGAAGGGGCCATCGAATTTGTGAGGGCCGAGCAGGAAA	900
QY	901 TAATGAAGAAGTTCCGGAGGGCTATGAGAATGATGTGGCTGCCATGAGCTCTCACTGA	960
DB	901 TAATGAAGAAGTTCCGGAGGGCTATGAGAATGATGTGGCTGCCATGAGCTCTCACTGA	960
QY	961 TCCCTGGCCTGAACCTGGCTGTAGGTCTTTCCAGCTTCATCCAGCGCAGTCCCGC	1020
DB	961 TCCCTGGCCTGAACCTGGCTGTAGGTCTTTCCAGCTTCATCCAGCGCAGTCCCGC	1020
QY	1021 CGCTCCAGCAGCGTTACTGGGGTGTCTCTATAGTCTCTTTATGAGGCTCCCGAGC	1080
DB	1021 CGCTCCAGCAGCGTTACTGGGGTGTCTCTATAGTCTCTTTATGAGGCTCCCGAGC	1080
QY	1081 AGGATGGTGCAGGTGTTTATCCCGCCAGGAGTGGCGCCATCATCGGCAAGAAAG	1140
DB	1081 AGGATGGTGCAGGTGTTTATCCCGCCAGGAGTGGCGCCATCATCGGCAAGAAAG	1140
QY	1141 GGCAGACATCAAAACAGCTCTCCCGGTTTGCAGCGCTTCCATCAAGATTGACACCGC	1200
DB	1141 GGCAGACATCAAAACAGCTCTCCCGGTTTGCAGCGCTTCCATCAAGATTGACACCGC	1200
QY	1201 AAACACCTGACTCCAAAGTTTATCGTATGGTTATCATCTGGACCGCAGAGGCCAATTC	1260
DB	1201 AAACACCTGACTCCAAAGTTTATCGTATGGTTATCATCTGGACCGCAGAGGCCAATTC	1260
QY	1261 AGGCTCAGGAAAGATCTATGCAAACTCAAGAGGAGAACTCTTTGGTCCCAAGAGG	1320
DB	1261 AGGCTCAGGAAAGATCTATGCAAACTCAAGAGGAGAACTCTTTGGTCCCAAGAGG	1320
QY	1321 AAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCTGCGCGGTCATTGGCA	1380
DB	1321 AAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCTGCGCGGTCATTGGCA	1380
QY	1381 AAGTGGAAAAACGGTGAACGAGTTTCAGAAATTCAGGAGGAGAACTCTTTGGTCCCAAGAGG	1440
DB	1381 AAGTGGAAAAACGGTGAACGAGTTTCAGAAATTCAGGAGGAGAACTCTTTGGTCCCAAGAGG	1440
QY	1441 GAGACAGACCCCTGATGAGAACGACGACCATCTGTGAAAATCATCGGACATTTCTATG	1500
DB	1441 GAGACAGACCCCTGATGAGAACGACGACCATCTGTGAAAATCATCGGACATTTCTATG	1500
QY	1501 CCAAGTGAAGTGGCTCAACGGAAGATCCGAGACATCTGGCCAGGTTAAGCAGAGCATC	1560
DB	1501 CCAAGTGAAGTGGCTCAACGGAAGATCCGAGACATCTGGCCAGGTTAAGCAGAGCATC	1560
QY	1561 AGAAGGACAGAGTAAACCGGCGCCAGGACGAGGAGTCCAGACGCGCTCCCTGTCCCT	1620
DB	1561 AGAAGGACAGAGTAAACCGGCGCCAGGACGAGGAGTCCAGACGCGCTCCCTGTCCCT	1620
QY	1621 TNGAGTCCAGACAAACGCGGAGAAATCGAGAGTGTCTCTCCCGGAGGCTCAG	1680
DB	1621 TNGAGTCCAGACAAACGCGGAGAAATCGAGAGTGTCTCTCCCGGAGGCTCAG	1680
QY	1681 ATGAGTGGGAATCCGGGACACNTGGCGGGCTGTAGATCAGGTTTGGCCACTTGATGA	1740
DB	1681 ATGAGTGGGAATCCGGGACACNTGGCGGGCTGTAGATCAGGTTTGGCCACTTGATGA	1740
QY	1741 GAAAGATTTCCAGTGAAGAACCTGATCTNTCAGCCCAACCAACCCCAATTTGGGCC	1800
DB	1741 GAAAGATTTCCAGTGAAGAACCTGATCTNTCAGCCCAACCAACCCCAATTTGGGCC	1800
QY	1801 AACACTGNTGCCCTCGGGGTGTGAGAAATTTAGCGCAAGGACATTTTAAACGTGGAT	1860
DB	1801 AACACTGNTGCCCTCGGGGTGTGAGAAATTTAGCGCAAGGACATTTTAAACGTGGAT	1860
QY	1861 TGTTTAAAGAAAGTCTCCAGGCCCAACCAAGAGGGTGGATCACACCTCAGTGGGAAGAA	1920
DB	1861 TGTTTAAAGAAAGTCTCCAGGCCCAACCAAGAGGGTGGATCACACCTCAGTGGGAAGAA	1920
QY	1921 AATTAATTTCTTCCAGGTTTTTAAAA 1946	
DB	1921 AATTAATTTCTTCCAGGTTTTTAAAA 1946	
XX	AAZ36151 standard; DNA; 1708 BP.	
AC	AAZ36151;	
RESULT 2		
AAZ36151		
ID	AAZ36151	
XX		
AC		

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2004, 03:34:49 ; Search time 719.632 Seconds
(without alignments)
11503.795 Million cell updates/sec

Title: US-09-270-437D-7
Perfect score: 1946
Sequence: 1 gctgtacggagggtctggg.....atttccttcagggttttaaaa 1946

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1941	99.7	1946	3	Aaz36153 An altern
2	1699.2	85.8	1708	3	Aaz36151 DNA encod
3	1389	71.4	2780	4	Aak94782 Human ful
4	1297.8	66.7	2224	2	Aaz10617 cDNA enco
5	1085.4	55.8	1182	4	Aas26148 Human cDN
6	1085.4	55.8	1182	7	Abx73489 Human nov
7	1069.8	55.0	1129	7	Aas26566 Human cDN
8	1069.8	55.0	1129	7	Abx73907 Human nov
9	732.8	37.7	833	4	Aak91969 Human cDN
10	732.8	37.7	833	4	Aak93655 Human cDN
11	674.2	34.6	1740	3	Aac66035 Human lun
12	674.2	34.6	1740	6	AbL49254 Human lun
13	674.2	34.6	1740	6	AbQ92440 Human lun
14	674.2	34.6	1743	6	AbL49299 Human lun
15	674.2	34.6	1743	6	AbL49297 Human lun
16	674.2	34.6	1743	6	AbQ92483 Human lun
17	674.2	34.6	1743	6	AbQ92485 Human lun
18	674.2	34.6	1743	8	Ada28540 Recombina
19	674.2	34.6	1743	8	Ada28537 Recombina
20	674.2	34.6	1764	6	AbL49283 Human lun
21	674.2	34.6	1764	6	AbQ92469 Human lun
22	674.2	34.6	1764	8	Ada28518 Recombina
23	672.6	34.6	1740	8	Ada28437 Human lun

24	672.6	34.6	4159	3	Aaz36150	DNA encod
25	672.6	34.6	4181	3	Aac65900	Human lun
26	672.6	34.6	4181	6	AbL65405	Lung canc
27	672.6	34.6	4181	6	AbL49119	Human lun
28	672.6	34.6	4181	6	AbQ92305	Human lun
29	672.6	34.6	4181	8	Ada28285	Human lun
30	672.6	34.6	4181	9	ADDI4661	Human lun
31	672.6	34.6	4181	9	ADBS3470	Human lun
32	672.6	34.6	4434	8	ACH03940	Human CDN
33	672.6	34.6	4601	6	ABA99958	Human KOC
34	620.8	31.9	1985	5	Aas76779	Human KOC
35	602.6	31.0	4264	5	Aas86150	DNA encod
36	585.4	30.1	3283	3	Aaz36154	DNA encod
37	585.4	30.1	3667	5	Aas70982	DNA encod
38	585.4	30.1	3667	6	ABs76442	cDNA enco
39	570.6	29.3	2010	5	Aas70981	DNA encod
40	570.6	29.3	3412	3	Aaz36152	DNA encod
41	570.6	29.3	3694	4	Aas26150	Human cDN
42	570.6	29.3	3694	7	Abx73491	Human nov
43	532.2	27.3	1707	7	ACA90176	cDNA enco
44	267.2	13.7	2290	4	Aas26153	Human CDN
45	267.2	13.7	2290	7	Abx73494	Human nov

ALIGNMENTS

RESULT 1

Aaz36153
ID Aaz36153 standard; DNA; 1946 BP.
XX
AC Aaz36153;
XX
DT 11-FEB-2000 (first entry)
XX
DE An alternative form of DNA encoding cancer associated antigen KOC-2.
XX
KW Cancer associated antigen; KOC-2; cancer; vaccine; CT7; ss.
XX
OS Homo sapiens.
XX
PN WO9954738-A1.
XX
PD 28-OCT-1999.
XX
PF 16-MAR-1999; 99WO-US005766.
XX
PR 17-APR-1998; 98US-00061709.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old LJ;
XX
DR WPI; 2000-013284/01.
XX
PT Nucleotides representing cancer-associated genes, used to develop
XX products for the diagnosis, monitoring and treatment of cancers.
XX
PS Claim 55; Page 42; 44pp; English.
XX
CC The present sequence represents an alternative form of a cancer
XX associated antigen gene designated KOC-2. The specification also
XX describes a cancer associated antigen designated CT7. The CT7
XX polynucleotide has been isolated from SK-MEL-37 melanoma cells. The
XX polynucleotide has some homology with MAGE-10, limited to about 210 carboxy
XX terminal amino acids. The amino terminal of the protein has a repetitive
XX pattern, with repeats rich in serine, proline, glutamine and leucine, and
XX an almost invariable core of the peptide given in AAY43877. The CT7
XX polypeptide can be processed to peptides which provoke lysis by cytolytic
XX T cells. The polynucleotides and polypeptides can be used for treating a
XX cancerous condition and screening for or diagnosing cancerous conditions.
XX The cancer associated antigens can be used as an immunogenic or vaccine
XX composition with an adjuvant, e.g. a cytokine, a saponin, or granulocyte

This Page Blank (uspto)

Search completed: July 24, 2004, 13:04:18
Job time : 7502.35 secs

Db 1777 GCCGAGGTGTCGTGCCGAGAGACCAGACCCAGACGAGAACGACCAAGTCATTGTTAAG 1836
QY 1482 ATCATCGGACATTCTTATGCCAGTCAGATGCTCAACGGAAGATCCGAGACATCCTTGCC 1541
Db 1837 ATCATCGGACATTCTTATGCCAGTCAGATGCTCAACGGAAGATCCGAGACATCCTTGCC 1896
QY 1542 CAGGTTAAGCAGCAGCATCAGAGGGACAGAGTAACAGCCAGCCAGGACCGGAGGAGTGA 1601
Db 1897 CAGGTTAAGCAGCAGCATCAGAGGGACAGAGTAACAGCCAGGACCGGAGGAGTGA 1956
QY 1602 -CCAGGCCCTCCCTGTCCTCCCTTNGAGTCAGGACAAACAGCGGCAGA----- 1646
Db 1957 CCGCGCCCTTCTCTGTCCTCCATGGCTCCAGATCAGCAGGCGGACACACAGGACTGGAGGAC 2016
QY 1647 -----ATCAGAGTGTGCTCTCCCGGAGGCTGAGATGAGTGGGAATCCGGAGAC 1701
Db 2017 GGTGAGGCGCGCGTGTCTCTCTAGCAGGCTGAGATGAGTGGGAATCCGGAGAC 2075
QY 1702 NTGGGCGGCGCTGTAGATCAGGTTTGCCCA 1731
Db 2076 TTGGGCTGCTGGCTGAGATCAGGTTTGACA 2105

RESULT 15
AF026527
LOCUS 2021 bp mRNA linear VRT 30-OCT-1997
DEFINITION Gallus gallus zipcode-binding protein (ZBP1) mRNA, complete cds.
ACCESSION AF026527
VERSION AF026527.1 GI:2570920
KEYWORDS
SOURCE
ORGANISM Gallus gallus (chicken)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 2021)
AUTHORS Ross,A.F., Oleynikov,Y., Kislauskis,E.H., Taneja,K.L. and Singer,R.H.
Characterization of a beta-actin mRNA zipcode-binding protein
Mol. Cell. Biol. 17 (4), 2158-2165 (1997)
97220007
121465
REFERENCE 2 (bases 1 to 2021)
AUTHORS Oleynikov,Y.S., Ross,A.F. and Singer,R.H.
Direct Submission
TITLE Submitted (24-SEP-1997) Cell Biology, Albert Einstein College of
JOURNAL Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, USA
FEATURES
Location/Qualifiers
1..2021
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/cell_type="fibroblast"
/dev_stage="embryo"
1..2021
/gene="ZBP1"
129..1859
/gene="ZBP1"
/function="binds a zipcode element in beta-actin mRNA
3'UTR and is involved in its localization"
/notes="beta-actin mRNA zipcode-binding protein"
/codon_start=1
/product="zipcode-binding protein"
/protein_id="AAB82295.1"
/db_xref="GI:2570921"
/translation="MNKLVIGNESVTPADLEKVFNDHKISFGSGLVKSGYAFVDC
PDEQWAKAIEFTSGVLEHKGQLEIHEHSVPKKQRKQIRNIPQLRWELDGLLA
QYGTWENCQVNTDSEAVNVYTNREOTROAIMKLNHOLENHLVKVSYIPDEQSV
QCPENRGSGFGARGAPRGSPVTAGAPVKQOPVDIPRLLVPTOVVGAIGKEGATI
RNITKQSKIDVHRKENAGAEKALISHTPEGCSAAKMILEIMQKEAKDTYADE
VELKILAHNFVGLRIGKGRNLRKVEQDTETKTISSIDQDLTYNPERTIIVKGSIE
NCKKAEQIMKKVREAVNDVAMSLQSHLIPGLNLAAYGLFPASSNAVPPPPSVSG
AAPYSFMPPEQETHVFIPAQAVGAILGKGQHIKQLSRFASASIKIAPPETPDSKV

misc_feature /note="encodes RNP2 RNA binding domain"
misc_feature /note="encodes RNP1 RNA binding domain"
misc_feature /note="encodes nuclear localization signal; NLS"
misc_feature /note="encodes nuclear localization signal; NLS"
misc_feature /note="encodes KH-domain #1"
misc_feature /note="encodes KH-domain #2"
misc_feature /note="encodes REV-like nuclear export signal; NES"
misc_feature /note="encodes KH-domain #3"
misc_feature /note="encodes KH-domain #4"
ORIGIN
Query Match 53.8%; Score 1046; DB 5; Length 2021;
Best Local Similarity 82.1%; Pred. No. 8.3e-246;
Matches 1231; Conservative 0; Mismatches 260; Indels 9; Gaps 2;
QY 108 AGGAGCGCGGAAATTCAAATCCGAAATATCCACCCAGCTCCGATGGGAAGTACTGGAC 167
Db 363 AGGAGCGCGGAAATTCAAATCCGAAATATCCACCCAGCTCCGATGGGAAGTACTGGAT 422
QY 168 AGCTCTGCTGCTCAGTATGTTAGTACAGTAGAGAACTGTGAGCAAGTGAACACCGAGAGTGAG 227
Db 423 GGCTTGCTAGCAATATGTCACCTGTAGAAACTGTGAGCAAGTGAACACAGACAGTGAAG 482
QY 228 ACAGCAGTGTGATGATCTACCTATTCACCGGAGACAGACAGGCAAGCCATCATGAAG 287
Db 483 ACCGCTGTGTTAATGTCTACCTACACAAACCGGAGACAGACAGGCAAGCCATCATGAAG 542
QY 288 CTGAATGGCCACAGTTGGAGAACCATGCCCTGAAGGTCTCTACATCCCGATGAGCAG 347
Db 543 TTAAACGGGACACAGCTGGAGAACCACTGCTGTGAAGTCTCTACATCCCTGATGAGCAG 602
QY 348 ATAGCAGAGGACCTGAGAAATGGCGCCGAGGGGGCTTTGGCTCTCGGGGTTCAGCCCGC 407
Db 603 TCCGTGACGGGCCAGAGAAATGGCGGGGGTGGCTTTGGGGCCCGAGGTGCCCGCCG 662
QY 408 CAGGGCTACCTGTGGCAGCGGGGGCCCGAGCCAGCAGCAGCAAGTGAACATCCCGCTT 467
Db 663 CAGGGCTCCCTGTCTCACTGAGGGGCTCCAGTCAAGCAGCAGCCCGTGGACATCCCGCTC 722
QY 468 CGGCTCTCTGGTCCCAACCCAGTATGTGGGTGCCATTTATTGGCAAGAGGGGGCCACCATC 527
Db 723 CGTCTGCTGGTCCCAACCGCAGTATGTGGGTGCCATCATTTGGCAAGAGGGGGCCACCATC 782
QY 528 CGCAACATCAAAAAAGACCCAGTCCAAGATAGACGTGTCATAGGAAGAGAGAACCGAGT 587
Db 783 AGGAACATTACCAAGCAGACAGTCCAAGATTGATGTGCAACCGTAAGSAGAAATGCAGGA 842
QY 588 GCAGCTGAAAAAGCCATCAGTGTGCACCTCCAGGGGTGCTCTCCCGCTGTGAAG 647
Db 843 GCTGCAGAAAAAGCTATCAGCATCCACTCCACCCCGGGGGTGTCTCTGCTGCCTGCAAA 902
QY 648 ATGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCAAAACCGCTGACGAGGTGCC 907

RESULT 14
AF541940 2105 bp mRNA linear ROD 25-NOV-2003
LOCUS Rattus norvegicus b-actin zipcode binding protein 1 mRNA, complete
DEFINITION cds.
ACCESSION AF541940
VERSION AF541940.1 GI:27464837
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 2105)
AUTHORS Eom, T., Antar, L.N., Singer, R.H. and Bassell, G.J.
TITLE Localization of a beta-actin messenger ribonucleoprotein complex
with zipcode-binding protein modulates the density of dendritic
filopodia and filopodial synapses
J. Neurosci. 23 (32), 10433-10444 (2003)
MEDLINE 22979172
PUBMED 14614102
REFERENCE 2 (bases 1 to 2105)
AUTHORS Eom, T., Singer, R.H. and Bassell, G.J.
TITLE Direct Submission
JOURNAL Submitted (27-AUG-2002) Neuroscience, Albert Einstein College of
Medicine, 1300 Morris Park Ave., Bronx, NY 10461, USA
FEATURES
source
1..2105
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
223..1956
/note="rZBP1"
/codon_start=1
/product="b-actin zipcode binding protein 1"
/protein_id="AA016210.1"
/db_xref="GI:27464838"
/translation="MNKLYIGNLNESTVPADLEKVPFAHKLSYSGQFLVKGYPVDC
PDEHWMKAIETFSQKVELOKLEIHSVPKORSRKIOIRNIPPOLRWBLVLSLLA
QYGVTVNCEQNTSETAVNVTYSNREQTQAIMKLNHGLEHNLKVSYPDEOLA
QYGVNRRGGSGQPRQSPVAGAPAKQQQVDIPURLVFPYQYVGAIIKKEGATI
RNITQKSDIVHRKENAGAAKAISSVHSPGSSACKMILEMHKEADKTADIE
VPLKILAHNNFVRLIIGXENLAKVQEDTKITISLQDLTYNPERPTIVKGAIE
NCCRAEQIIMKKVREAVENDVAASLQSHLIPGLNLAAGLFPASSSVAPPPSVTG
AAPYGSFWQAPQEMVOVFIQAQAVGALIGKKGOHIKOLSRFASAIKIAPETPDSK
VRMVITGPFAQFAQRIYIGKLEENFPFGKEVLETHIRVPASAGSVIGKGGK
TVNELQNLTAAEVVVPRDQTPDENDQVIVKLIIGHFYASQMAQRKIRDLIAQVKQHQK
GQSNQAQARRK"
ORIGIN
Query Match 64.2%; Score 1250; DB 10; Length 2105;
Best Local Similarity 86.7%; Fred. No. 6.4e-296;
Matches 1431; Conservative 0; Mismatches 192; Indels 27; Gaps 4;
QY 108 AGGAGCGCGGAAATTCAAATCCGAATATTCACCCAGCTCCGATGGGAAGTACTGGAC 167
DB 457 AGGAGTCGGAAATACAGTCCGGAATATTCACCTCAGCTCCGATGGGAAGTCTGGAT 516
QY 168 AGCCTGCTGGCTCAGTATGGTACAGTAGAGAACTGTGAGCAAGTGAACACCCAGAGTGAG 227
DB 517 AGCCTTCTGGCCAGTACGGCAGTGGAGAACTGTGAGCAAGTGAACACCCAGAGTGAG 576
QY 228 ACGGAGTGTGAATGTACCTATTCCACCGGGAGCAGACCCAGGCAAGCCATCATGAAG 287
DB 577 ACGGCGGTGGTCAACGCTACCTATCTAACCGGGAGCAGACCCAGGCAAGCCATCATGAAG 636
QY 288 CTGAATGGCCACAGTTGGAGAACCATGCCCTTGAAAGGTTCTCTACATCCCCGATGAGCAG 347
DB 637 CTAATGGCCATCAACTGGAGAACCATGCCCTTGAAAGGTTCTCTACATCTCATGAGCAG 696
QY 348 ATAGCACAGGACCTGAGAATGGGCGCGAGGGGGCTTTGGCTCTCGGGGTGAGCCCCG 407

Db 697 ATAGCACAGGTCCTGAGATGGGCGCGCTTGGAGGCTTTGGTCTCGAGGCGAGCCAGG 756
QY 408 CAGGGCTCACCTGTGGCAGCGGGGGCCCCAGCCAGCAGCAGCAAGTGGAGCATCCCCCTT 467
Db 757 CAGGGTCAACGGTAGCAGCAGGAGGGCGCCAGCCAGCAGCAGCAGTGGAGCATCCCTCTC 816
QY 468 CGGCTCCTGTGCCACCCAGTATGTGGGTGTCCTATTATTCGACAGGAGGGGCCACCATC 527
Db 817 CGGCTCCTGTGCCACCCAGTATGTAGGCGCCATCATTTGGCAAGAGGGGTGCCACCATC 876
QY 528 CGCAACATCAAAACAGACCCAGTCCAAAGATAGACGTGTCATAGGAAGAGAGACGACGT 587
Db 877 CGAAACATCAAAACAGACTCAGTCCAGATAGACGTGTCACAGGAAGAGAGACGCGGG 936
QY 588 GCAGCTGAAAAGGCATCAGTGTGCATCCACCCTGAGGCTGTCTCTCGCTTGTGAAG 647
Db 937 GCTCGGAGAGAGCCATCAGTGTGCATCCGACCCCTGAAGGCTGTCTCTCGGCTGTGAAG 996
QY 648 ATGATCTTGGAGATTATGATTAAGAGGCTAAGGACACCAAAACGGCTGACGAGTTCCT 707
Db 997 ATGATCTTGGAGATTATGCAACAGGAGCAAAAGGACACCAAGACGGCAGACGAGATTCC 1056
QY 708 CTGAAGATCTGGCCCATATAAATCTTGTAGGGGCTCTCATTTGGCAAGAGAGACGGAAC 767
Db 1057 CTGAAGATCTTGGCTCATATAAATCTTGTGGGGGACCTCATTTGGAAGAGAGGGCGGAAC 1116
QY 768 CTGAAGAGGTAGAGCAAGATACCGAGACAAAATACACATCTCTCTGTTGGAAGACCTT 827
Db 1117 TTGAAGAAAGTGGAGCAGATACAGAGACTAAGATCACCATCTCTCTCGCTCCAGGATCTC 1176
QY 828 ACCCTTCAACCCCTGAGAGACCATCACTGTGAGGGGGCCATCGAGAAATTTGTCAGG 887
Db 1177 ACGCTCTATAACCCCTGAAAGACCATATACTGTGAAGGGGGCCATTTGAGAATCTGTGCG 1236
QY 888 GCCGAGCAGGAAATAATGAAGAAAGTTCGGGAGGCTATGAGAAATGATGTGCTGCCATG 947
Db 1237 GCTGAGCAGAGATCATGAGAAAGTTCGAGAGGCTTACGAGAAATGACGTGCTGCCATG 1296
QY 948 AGC-----TCTCACTCTGATCCCTGGCTGAACTCTGGCTGTGTAGTGTCTTTTCCAGCT 1001
Db 1297 AGCTTGCAGTCTCACTCTCATCTCTGCTCACTCTGCTGCTGCTGCTCTCTTCCAGCT 1356
QY 1002 TCATCCAGCAGTCCCGCCGCTCCGAGCAGCTTACTGGGCTGCTCCCTATAGTCTCC 1061
Db 1357 TCTTCTAGTGGCTTCCGCTCTCTCCAGCAGCTCCTCTGGGGCTGCTCCCTATGCTCC 1416
QY 1062 TTTATGCAAGTCCCGAGCAGGAGATGCTGCAAGTGTATTATCCCGCCCGCAGGAGTGGC 1121
Db 1417 TTTATGCAAGTCCCGAGCAGGAGATGTTACAGTGTTCATCCCGCCCGCAGGCTGTGGC 1476
QY 1122 GCCATCATCGGCAAGAGGGCAGCAGATCAAAACAGCTCTCCCGTTTGGCAGCGCTCC 1181
Db 1477 GCCATCATCGGCAAGAGGGCCAGCAGATCAAAACAGCTCTCCCGTTTGGCAGCGCTCC 1536
QY 1182 ATCAAGATTGACCCCGAAGACCTGACTCCAAAGTTCGATGCTGTTATCTACTCTGGA 1241
Db 1537 ATCAAGATTGCTCCACCAAGAACACCTGACTCCAAAGTTCGAAATGTTGTCTACTGGA 1596
QY 1242 CGCCAGAGGGCCCAATTCAAGGCTCAGGAGAAATCTATGGCAAACTCAAGGAGAGAAC 1301
Db 1597 CCCCAGAGGCTCAGTTCAAGGCTCAGGAGAAATCTATGGCAAACTAAAGAGAGAAAT 1656
QY 1302 TTTCTTTGTTCCAAAGAGAGTGAAGCTGAGAGACCCATACATAGTGTGCCAGCATCAGCA 1361
Db 1657 TTTCTTTGTTCCAAAGAGAGTGAAGCTGAGAGACCCATCCGGGTCCAGCTTTCAGCA 1716
QY 1362 GCTGCCCGGCTCATTTGGCAAGGTCGAAACCGGTGAACGAGTTCAGAAATTTGACGCA 1421
Db 1717 GCTGCCCGGCTCATTTGGCAAGGTCGAAACCGGTGAACGAGTTCAGAAATTTGACGCA 1776
QY 1422 GCTGAGGTGGTAGTACCAGAGACCCCTGATGAGAACGACGAGTCTATCGTGA 1481

QY 1302 TTCTTTGGTCCCAAGAGGAAAGTGAAGCTGGAGACCCACATACGTGTGGCAGCATCAGCA 1361
Db 1565 TTCTTTGGTCCCAAGAGGAAAGTGAAGCTGGAGACCCACATACGTGTGGCAGCATCAGCA 1624
QY 1362 GCTGGCCGGGTCTATTGCAAGGTGGAAAGCGGTGAACGAGTGTGCAGAAATTTGACGCGCA 1421
Db 1625 GCGGGCCGGGTCTATTGCAAGGTGGAAAGCGGTGAACGAGTGTGCAGAAATTTGACGCGCA 1684
QY 1422 GCTGAGGTGGTGTAGTACCAAGAGACCAACCCCTGATGAGAACGACCAACGAGTGTGATCGTGA 1481
Db 1685 GCTGAGGTGGTGTAGTACCAAGAGACCAACCCCTGATGAGAACGACCAACGAGTGTGATCGTGA 1744
QY 1482 ATCATCGGACATTTCTATGCGAGTCAAGTGGCTCAACGAGAGATCCGAGACATCTCGGCC 1541
Db 1745 ATCATCGGACATTTCTATGCGAGTCAAGTGGCTCAACGAGAGATCCGAGACATCTCGGCC 1804
QY 1542 CAGGTTAAGCAGCAGCATCAGAGGGACAGAGTAAACAGGCCACCGAGGACGAGGAGAGTGA 1601
Db 1805 CAAAGTTAAGCAACAGCAGCAGAGGGACAGAGTAAACAGGCCACCGAGGACGAGGAGAGTGA 1864
QY 1602 -CCAGCCCTCCCTGCTCCCTTNGAGTCCAGGACCAACCAACGCGGACAGAA----- 1647
Db 1865 CCGCGCCCTCCCTGCTCCCTTNGAGTCCAGGACCAACCAACGCGGACAGAA----- 1924
QY 1648 -----ATCGAGAGTGTGCTCTCCCGCGCAGCCTGAGAAATGAGTGGGAATCCGGGAC 1699
Db 1925 GCGGGTGGAGGGCGCGGTGTGTTTTTCCAGCAGGCTGAGAAATGAGTGGGAATCCGGGAC 1983
QY 1700 ACNTGGCGCGGCTGTAGATCAGGTTTCCCACTGTGTTGAGAAAGATGTTCCAGTGGG 1759
Db 1984 ATTTGGGCGCTGGTGGAGATCAGGTTTGCACACTGTATTGAGAAACAATGTTCCAGTGGG 2043
QY 1760 AACCTGATCTNTCAGCGCCCAACACCCCACTATGGCCCAACCACTGTNTGCCCTCGG 1819
Db 2044 ATCTGATCTCTGCCCCCAA--TTGAGCCAGCTGGCCACAGCCCTTGGATAT 2101
QY 1820 GGTGTCAGAAATNTAGCGAAGGACATTTTAAACGTGGATGTTTAAAGAGCTCTCCA 1879
Db 2102 CACCATTTGCAATCATAGCTGGTGTCTTTTAAACGTGGATGTTCT--TGAAGTTCTCCA 2159
QY 1880 GGCCCCCACCAGAGGGTGGATCACCTCAGTGGGAGAGAAATAAATTTCCCTTCAGT 1939
Db 2160 GCCTCCATGGA-AGGATGGGTGAGATCCCATCCAGTGGGAGAGAAATAAATTTCCCTTCAGT 2218
QY 1940 TTTA 1943
Db 2219 TTTA 2222
RESULT 13
LOCUS BC051679 2444 bp mRNA linear ROD 08-OCT-2003
DEFINITION Mus musculus insulin-like growth factor 2, binding protein 1, mRNA
(cDNA clone MGC:60613 IMAGE:30008106), complete cds.
ACCESSION BC051679
VERSION BC051679.1 GI:30354043
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Strausberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raba, S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Vallalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahney, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A.C., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalilus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
PUBMED 12477932
REFERENCE 2 (bases 1 to 2444)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (30-APR-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Kenneth R. Boheler (National Insitute on
Aging, USA)
cDNA Library Preparation: Yulan Piao and Minoru Ko (National
Institute on Aging, NIH: http://lgsun.grc.nia.nih.gov/cDNA/)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amgobcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Series: IRAC Plate: 111 Row: e Column: 24
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6753517.

FEATURES
source

1..2444
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129/Sv x 129/Sv-CP"
/db_xref="taxon:10090"
/clone="MGC:60613 IMAGE:30008106"
/tissue_type="ES Cell, undifferentiated, mouse, R1
Embryonic stem cells"
/clone_lib="NIA Mouse Undifferentiated ES Cell cDNA
Library (Long)"
/lab_host="DH10B"
/note="Vector: pSPORT1"
1..2444
/gene="Igfbpl"
/note="synonyms: IMP-1, Neilsen, CRD-BP"
/db_xref="LocusID:140486"
/db_xref="MGI:1890357"
295..2028
/codon_start=1
/product="insulin-like growth factor 2, binding protein 1"
/protein_id="AAH51679.1"
/db_xref="GI:30354044"
/db_xref="LocusID:140486"
/translation="MNKLYIGNLNESTVPADLEKVFABHKISYSGQFLVKGAFVDC
PDEHWAKAIETFSKVLQKLEIHSVPKORSRKIQIRNPQQRWEVLDSLLA
QYGTVEQVNTSEAVNVTVSNRQTAQKLNHOLENHALKXVYIPDQIT
QCPNRRGGFGSGQPGQKSPVAGAPAKQPDVPIRLVPTVYGAIIKGGAT
RNIKTQSKIDVHRKNAGAAKSAISVHSTPEGSSACKMILEIMHKEADTKTAG
VPLKLAHNFFVGLIGIKRNLKVEQDTETKTISSQLDTLTFNPTTIPVGAIE
NCCRAEQIMKKVREAYENDVAASQSLHLPGLNLAAGFPASSSSAVPPPPVTG
AAPYSFMQAPEQEMVQVFIIPAQAVGAIIGKQGHKQLSRFASIKIAPPETPDSK

gene

CDS

REFERENCE	2 (bases 1 to 2223)	QY	228	ACGCGAGTGGTGAATGTCACCTATTCCAAACCGGAGCAGACACGAGCAAGCATCATGAAG	287
AUTHORS	Herrick,D.J. and Ross,J.				
TITLE	The half-life of c-myc mRNA in growing and serum-stimulated cells: influence of the coding and 3' untranslated regions and role of ribosome translocation	Db	485	ACAGCGTGGTCAACGCTCACTACTCTAAACCGGAGCAGACGAGCAAGCTATCATGAAG	544
JOURNAL	Mol. Cell. Biol. 14 (3), 2119-2128 (1994)				
MEDLINE	94159886	QY	288	CTGAATGGCCACCAAGTTGGAGACCAATGCGCTGAAGGTCTCTACATCCCGATGAGCAG	347
PUBMED	8114742	Db	545	CTAAATGGCCATCACTGGAGAACCATGCCCTGAAGGTCTCTACATACCTGATGAGCAG	604
REFERENCE	3 (bases 1 to 2223)				
AUTHORS	Prokipcak,R.D., Herrick,D.J. and Ross,J.	QY	348	ATAGCACAGGGACCTCTAGAAATGGGCGCCGAGGGGGCTTTGGCTCTCGGGGTCAAGCCCGC	407
TITLE	Purification and properties of a protein that binds to the C-terminal coding region of human c-myc mRNA	Db	605	ATACACACAGGTCTCTGAGATGGGCGCTCTGAGAGGTCTTTGGGTCTCGGGGCGAGCCCGG	664
JOURNAL	J. Biol. Chem. 269 (12), 9261-9269 (1994)				
MEDLINE	94179348	QY	408	CAGGGCTCACTGTGGCAGCGGGGGCCAGCCCAAGCAGCAGCAAGTGAACATCCCCCTT	467
PUBMED	8132563	Db	665	CAAGGGTCGCGGTGGCAGCAGGGGCTCCAGCAAGCAGCAGCGCATGGACATCCCTCTC	724
REFERENCE	4 (bases 1 to 2223)				
AUTHORS	Leeds,P., Kren,B.T., Boylan,J.M., Betz,N.A., Steer,C.J., Gruppiso,P.A. and Ross,J.	QY	468	CGGCTCCTGTGCCACCCAGTATGTGGTGCATTTATTGGCAAGGAGGGGCGCACATC	527
TITLE	Developmental regulation of CRD-BP, an RNA-binding protein that stabilizes c-myc mRNA in vitro	Db	725	CGGCTCCTGTGCTACGCAAGTATGTAGGCGCTATCAITGGCAAGGAGGTGCCACATC	784
JOURNAL	Oncogene 14 (11), 1279-1286 (1997)				
MEDLINE	97322234	QY	528	CGCAATCATCAAAAACAGACCCAGTCCAAAGATAGACGTGCATAGGAGAGAGAACGCGGT	587
PUBMED	9178988	Db	785	CGAAATCATCAAAAACAGACCGAGTCCAAAATAGACGTGCATAGGAGAGAGATCGGGC	844
REFERENCE	5 (bases 1 to 2223)				
AUTHORS	Doyle,G.A., Betz,N.A., Leeds,P.F., Fleisig,A.J., Prokipcak,R.D. and Ross,J.	QY	588	GCAGCTGAAAAAGCATCATGTGTGACTCCACCCCTGAGGGGTGCTCTCCGTTGTAAG	647
TITLE	The c-myc coding region determinant-binding protein: a member of a family of KH domain RNA-binding proteins	Db	845	GCTGCGGAGAAGCCATCAGCGTGCATTCACCCCTGAAGGTGCTCTCCGCGTCAAG	904
JOURNAL	Nucleic Acids Res. 26 (22), 5036-5044 (1998)				
MEDLINE	99030504	QY	648	ATGATCTTCGAGATTATGCATAAAGAGGTAAAGGACCAACCAACGCTGACAGAGTTCC	707
PUBMED	9801297	Db	905	ATGATCTTCGAGATTATGCACAAAGGAGGCAAGGACCAACCAACGCTGACAGTTCC	964
REFERENCE	6 (bases 1 to 2223)				
AUTHORS	Ross,J., Prokipcak,R.D., Leeds,P., Doyle,G.A.R., Betz,N.A. and Fleisig,A.J.	QY	708	CTGAAGATCCTGGCCCATTAATACTTTGAGGGCGTCTCAITGGCAAGAGAGACGGAAC	767
TITLE	Direct Subission	Db	965	CTGAAGATCCTGGCTCATAACAACCTTCGTCGGCGACTCATTTGGCAAGAGAGCGCGAAC	1024
JOURNAL	Submitted (25-APR-1998) McArdle Laboratory, University of Wisconsin - Madison, 1400 University Ave., Madison, WI 53706, USA				
FEATURES	Location/Qualifiers	QY	768	CTGAAGAGGTAGAGCAAGATACCGAGACAAAATCACCATCTCTCTGTTGCAAGACCTT	827
source	1. .2223	Db	1025	CTGAAGAGGTGGAGCAGGACACAGAGACGAAGATCACCATCTCATCGCTCCAGGACCTC	1084
	/organism="Mus musculus"				
	/mol_type="mRNA"	QY	828	ACCTTTTCAACCTCAGAGGACCATCACTGTGAAGGGGCGCATCAGAGATTTGTTCCAGG	887
CDS	/db_xref="taxon:10090"	Db	1085	AGCTCTATAACCTGAGAGGACCATCACTGTGAAGGGCGCCATTGAGAACTGTTGCAGG	1144
	131..1864				
	/function="putative c-myc mRNA binding protein"	QY	888	GCCGAGCAGGAATAATGAAGAAAGTTCCGGAGGCGCTATGAGAATGATGTGGCTGCATG	947
	/notes="CRD-BP"	Db	1145	GCCGAGCAGGAGATCATGAAGAAAGTTCCGAGAGGCTTACGAGAACGACGTGGCGCATG	1204
	/codon_start=1				
	/product="coding region determinant binding protein"	QY	948	AGC-----TCTCACTGATCCCTGCGCTGAACCTGGCTGCTGTAGTCTTTTCCAGCT	1001
	/protein_id="AAC72743.1"	Db	1205	AGCTTGCAGTCCCACTCATCTCCCTGGGCTTAACTGGCTGCTGTAGTCTCTTCCAGCT	1264
	/db_xref="GI:3273749"	QY	1002	TATCCAGCGCAGTCCCGCGGCTCCAGCAGCGTTACTGGGGCTCTCCCTATAGTCC	1061
	/translation="MNKLVLGNLSEVTADLEKVFAEHKISYSGOFLVKSGYAFVDC	Db	1265	TATCCAGCGCTGTCCTCTCTCTCCACAGTGTCACTGGGGCTGCTCCCTATAGTCC	1324
	PEHWAMKALETPSGKVELQGRLEIEHSVPKQSRKTIQIRNIPQLRWEVLDSLLA				
	QVGTVENCGVNTSESTAVVNVTSNREQTRQIMKNGHLENHAKLVSYIPDEGIT				
	QGENRRGFGSRGQPSGVPAGAPAKQQVDPIRLPLTVQTVGAIIGKEGATI				
	RNITKQSDIVHRKENAAGAKAISVHSTPEGSSACKMILEIMHKEAKDTKDAE				
	VPLILAHNFVGRLLTKEGRNLKVEQDTETKITISLQDLTYNPERTITVKGAIE				
	NCKRAQETMKVREAVENDVAMSLQSHLPLGLNLAAVGLPLPASSAVPPPSVTG				
	AAPYSFPMQAPQEMVQVFTPAQAVCAIIGKKQHLKQLSRFASASIKIAPETPSK				
	TVNVLITGPEAGVFKAGRIYGLKEENFFGPKVEVKLETHIRVPASAGRVIGKGGK	QY	1062	TTTATGCAAGCTCCCGAGCAGGAGATGGTGCAGGTTTATCCCGCCAGGAGGAGTGGG	1121
	YNNLQNLTAAEVVPVPRDTPDENDQVIVKIIGHFYASQMAQRKIRDILAQVKQHQK	Db	1325	TTTATGCAAGCTCCGAGCAGGAGATGGTACAAGTGTTCATCCCGCCCGAGGCTGTGGG	1384
	GQSNTLAQRKK"				
ORIGIN		QY	1122	GGCATCATCGCAAGNAGGGGAGACATCAACAGCTCTCCGGTTTCCGAGCGGCTCC	1181
Query Match	66.7%; Score 1298.6; DB 10; Length 2223;	Db	1385	GGCATCATTTGGCAAGAGGGCCAGCACATCAACAACTCTCCCGTTTCCGAGCGGCTCC	1444
Best Local Similarity	83.7%; Pred. No. 7.4e-308;				
Matches 1561; Conservative	0; Mismatches 269; Indels 34; Gaps 7;	QY	1182	ATCAAGATTGCACCAACCCGAAAACACTGACTCCAAAGTTTCGATGTTATCATCTAGTA	1241
QY	108	AGGAGCGGGAATAATCAATCCGAATATTCACACCCAGCTCCGATGGGAAGTACTGGAC	167		
Db	365	AGGAGTCGGAATAACAGATCCCAATATTCACCTCAGCTCCGATGGGAGTCTAGAT	424		
QY	168	AGCTGCTGGCTCAGTATGTAAGTAGAGAACTGTGAGCAAGTGAACACCGAGAGTGA	227		
Db	425	AGCTGCTGGCTCAGTACGCTGAGTGAAGAACTGTGAGCAAGTGAACACTGAAAGTGA	484		

Db 365 AGGAGTCGGAATAATACAGATCCGCAATATTCCACCTCAGCTCCGATGGGAAGTCTAGAT 424
Qy 168 AGCCTGCTGGCTCAGTATGGTACAGTAGAGAACTGTGAGCAAGTGAACACACGAGAGTGAG 227
Db 425 AGCCTGCTGGCTCAGTATGGTACAGTAGAGAACTGTGAGCAAGTGAACACACTGAAAGTGAG 484
Qy 228 ACGCAGTGTGATGTCACCTATTCCAAACGGGAGCGAGACACAGGCAAGCCATCATGAAG 287
Db 485 ACAGCGTGTGTCAACGTCACCTACTTAACCGGGAGCAGACAGGCAAGCTTATCATGAAG 544
Qy 288 CTGAATGGCCACAGTGTGGAAACCATGCTCCCTGAGGTCTCTACATCCCGGATGAGCAG 347
Db 545 CTAAATGGCCATCAACTGGAGAAACATGCTCCCTGAGGTCTCTACATCACTGATGAGCAG 604
Qy 348 ATAGCAGAGGACCTGGAATGSGCGCGAGGGGCTTTTGGCTCTCGGGGTGAGCCCGC 407
Db 605 ATACACAAGTCTGAGAAATGGGCGTCTGGAGGCTTTGGGTCTCGGGGCGAGCCCGG 664
Qy 408 CAGGGCTCACTGTGGCAGCGGGGGCCCCAGCCAGCAGCAGCAGTGGGACATCCCGCTT 467
Db 665 CAAGGTCGCGCGTGGCAGCAGGGGCTCCAGCAAGCAGCAGCAGTGGGACATCCCTCTC 724
Qy 468 CGGCTCCTGTGGTCCCAACCCAGTATGTGGTGCCTATTATTGCAAGGAGGGGCCACCATC 527
Db 725 CGGCTCCTGTGGTCCCAACCCAGTATGTGGTGCCTATTATTGCAAGGAGGGGTGCCACCATC 784
Qy 528 CGCAACATCACAAACAGACCCAGTCCAAAGATAGACGTGTCATAGGAAGAGAAACGACGT 587
Db 785 CGAAACATCACAAACAGACCCAGTCCAAATAGACGTGTCATAGGAAGAGAAATGCGGGC 844
Qy 588 GCAGCTGAAAAGCATCAGTGTGCACTCCACCCCTGAGGGTGTCTCTCGCTTTGTAAG 647
Db 845 GCTCGGAAGAGGCCATCAGCGTGCATTTCAACCCCTGAAGGTGTCTCTCGCGGTGCAAG 904
Qy 648 ATGATCTTGGAGATTATGCAATAAGAGGCTAAGGACACCAAAACGGCTGAGAGTCTCC 707
Db 905 ATGATCTTGGAGATTATGCAAAAGAGGCAAGGACACCAAAACGGCAGATGAAGTCTCC 964
Qy 708 CTGAAGATCTTGGGCCATAATAACTTTTGTAGGGCTCTCAITGGCGAGGAGGACGGAAC 767
Db 965 CTGAAGATCTTGGCTATAACAACTTCTGCGGGGACTCAITGGCAAGGAGGCGGAAC 1024
Qy 768 CTGAAGAGGTGAGGAGATACGAGACAAAAATCACCATCTCTCTGTTGCAAGACCTT 827
Db 1025 CTGAAGAGGTGAGGAGACACAGACAGAGATCAACATCTCATCTGCTCCAGGACCTC 1084
Qy 828 ACCCTTTACACCTTGAGGACCATCACTGTGAAGGGGCCATCGAGAAATTTGTTGACGG 887
Db 1085 ACGCTTATAACCTTGAGGACCATCACTGTGAAGGGGCCATCGAGAAATTTGTTGACGG 1144
Qy 888 GCCGAGCAGGAATAATGAAGAAAGTTCGGAGGCGCTATGAGAAATGATGTGCTGCCATG 947
Db 1145 GCCGAGCAGGAGATCATGAAGAAAGTTCGAGAGGCTTACGAGAACGAGTGGCGCCCATG 1204
Qy 948 AGC-----TCTACCTGATCTGGCTGAAACCTGGCTGTGTAGGTCTTTTCCAGCT 1001
Db 1205 AGCTTGCAAGTCCCACTCATCCCTGGCTTAACTGGCTGTGTAGGTCTCTTCCAGCT 1264
Qy 1002 TCATCCAGCGAGTCCCGCCTCCAGCAGCGTTACTGGGGCTGCTCCCTATAGTCC 1061
Db 1265 TCATCCAGCGTGTCCCTCTCCAGCGAGTGTCACTGGGGCTGCTCCCTATAGTCTC 1324
Qy 1062 TTTATGAGGCTCCCGAGCAGGATGTTGCGAGGTGTTTATCCCGCCCGAGCGAGTGGC 1121
Db 1325 TTATGAGGCTCCCGAGCAGGATGTTGCAAGTGTTCATCCCGCCCGAGCGTGGGCG 1384
Qy 1122 GCCATCATCGAAGAGGGGAGCAGATCAAGAGCTCTCCCGTTTGGCAGCGCTCC 1181
Db 1385 GCCATCATCGAAGAGGGGAGCAGATCAAGAGCTCTCCCGTTTGGCAGCGCTCC 1444
Qy 1182 ATCAAGATTGACCAACCCGAAACACTGACTCTCAAGTGTGATGTTATCATCACTGGA 1241
Db 1445 ATCAAGATTGCTCCACCAAGAAACACTGACTCTCAAGTGTGATGTTGCTCATCTGGA 1504

Qy 1242 CCGCCAGAGGGCCCAATTCAAGGCTCAGGGAAGAAATCTATGCAAACTCAAGGAGAGAAC 1301
Db 1505 CCCCAGAGGCTCAGTTCAAGGCTCAGGGAAGAAATTTATGGCAAACTAAAAGAGAGAA 1564
Qy 1302 TTCTTTGGTCCCAAGGAGGAGTGAAGCTGAGAGCCCAATACATGCTGTGCCAGCATCAGA 1361
Db 1565 TTCTTTGGTCCCAAGGAGGAGTGAAGCTAGAGACCCCAATACATGCGGTTCCGGCTTCAGA 1624
Qy 1362 GCTGGCGGGGTCAATTGGCAAAAGGTGGAACCGGTGAACGGTGTGCAAAATTTGACGCA 1421
Db 1625 GCCGGCGGCTCATCGCAAGGCGCAAAACGGTGAATGAGTGCAGAACTTGACTGCA 1684
Qy 1422 GCTGAGGTGTGATGACCAAGAGACAGACCCCTGATGAGAACGACCAAGTCACTCTGAAA 1481
Db 1685 GCTGAGGTGTGATGACCAAGAGACAGACCCCGGATGAGAACGACCAAGTCACTCTTAAG 1744
Qy 1482 ATCATCCGACATTTCTATGCCCAGTCAGATGCTCAACGGAAGATCCGAGACATCTGGCC 1541
Db 1745 ATCATCCGACATTTCTATGCCAGCCAGATGCTCAGCGAAGATCCGAGACATCTGGCT 1804
Qy 1542 CAGGTTAAGCAGCAGCATCAGAGGACAGAGTAAACAGGGCCAGGACGAGGAAAGTGA 1601
Db 1805 CAAGTTAAGCAACAGCACCAAGGACAGAGCAACCTGGCCAGGCAACGAGGAAAGTGA 1864
Qy 1602 -CCAGCCCTCCCTGTCTCCCTTNGAGTCCAGGACACAAACGGGACAA----- 1647
Db 1865 CCGCGCCCTCTCTGTCCCATTTGGCTCCAGATCAGCAGGAGGAAACAGAACTCGAGGG 1924
Qy 1648 -----ATCGAGAGTGTCTCTCCCGGAGGCTCAGAAATGAGTGGAAATCCGGAC 1699
Db 1925 CGGGTGAAGGCGCGTGTGTTTTCCAGCAGGCTCAGAAATGAGTGGAAATCCAGG-C 1983
Qy 1700 ACNTGGCGGGCTGTAGATCAGGTTTGCACATTTGATGAGAAAGATGTTCCAGTGAGG 1759
Db 1984 ATTTGGGCTGCTGGAGATCAGGTTTGCACACTGTAATTGAGAACAAATGTTCCAGTGAGG 2043
Qy 1760 AACCTTGATCTNTAGCCCCCAACACCCCAATGTCGCCCACTGNTGCCCTCGG 1819
Db 2044 AATCTGATCTCTCGCCCCCAA--TTGAGCCAGCTGGCCACAGCCCAACCTTGGAAAT 2101
Qy 1820 GGTGTCAAAATTTAGCGCAAGGCACTTTAAACGTGGATGTTTAAAGAGCTCTCCA 1879
Db 2102 CACATTTCAATCATAGCTTGGTGTCTTTAAAGCTGGAATGCT--TGAAGTTCTCA 2159
Qy 1880 GCGCCCAACCAAGAGGTTGATCACAACCTCAGTGGGAAGAAATAAATTTCTTCAGGT 1939
Db 2160 GCCTCCATGGAAGGATGGTCCAGATCCAGTGGGGAAGAGAAATAAATTTCTTCAGGT 2219
Qy 1940 TTTA 1943
Db 2220 TTTA 2223

RESULT 12

AF061569 2223 bp mRNA linear ROD 11-NOV-1998
LOCUS Mus musculus coding region determinant binding protein mRNA,
DEFINITION complete cds.
ACCESSION AF061569
VERSION AF061569.1 GI:3273748
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2223)
AUTHORS Bernstein, P.L., Herrick, D.J., Prokipcak, R.D. and Ross, J.
TITLE Control of c-myc mRNA half-life in vitro by a protein capable of
binding to a coding region stability determinant
JOURNAL Genes Dev. 6 (4), 642-654 (1992)
MEDLINE 92217743
PUBMED 1559612

Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center, cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).

FEATURES

source
1. .2780
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP3000789"
/cell_line="NT2"
/cell_type="Teratocarcinoma"
/clone_lib="NT2RP3"
/note="Cloning vector: pME18SFL3-mRNA from NT2 neuronal
precursor cells after 2-weeks retinoic acid (RA)
induction"

ORIGIN

Query Match 71.4%; Score 1389; DB 9; Length 2780;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1405; Conservative 0; Mismatches 5; Indels 6; Gaps 1;
537 ACAAACAGACCCAGTCCAAAGTAGAGCTGATAGGAGAGAACGCGAGTGCAGCTGAA 596
Db 1 ACAAACAGACCCAGTCCAAAGTAGAGCTGATAGGAGAGAACGCGAGTGCAGCTGAA 60
597 AAGCCATCAGTGTGACCTCCACCCCTGAGGCTGCTCTCCCTTGTAAAGATGATCTTG 656
Db 61 AAGCCATCAGTGTGACCTCCACCCCTGAGGCTGCTCTCCCTTGTAAAGATGATCTTG 120
657 GAGATTATGATGAAGAGGCTAAGGACCAACAAACGGCTGACGAGGTTCCCTCTGAAGATC 716
Db 121 GAGATTATGATGAAGAGGCTAAGGACCAACAAACGGCTGACGAGGTTCCCTCTGAAGATC 180
717 CTGGCCCAATAAATCTTTGTAGGCGCTCTCATTTGGCAAGGAGGACGGAACCTGAAGAAG 776
Db 181 CTGGCCCAATAAATCTTTGTAGGCGCTCTCATTTGGCAAGGAGGACGGAACCTGAAGAAG 240
777 GTAGAGCAGATACCGAGACAAATACCATCTCTCTCGTTGCAAGACCTTACCTTTTAC 836
Db 241 GTAGAGCAGATACCGAGACAAATACCATCTCTCTCGTTGCAAGACCTTACCTTTTAC 300
837 AACCTCTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAATTTGTTGACGGGCGGAGCAG 896
Db 301 AACCTCTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAATTTGTTGACGGGCGGAGCAG 360
897 GAATTAATGAAGAAAGTTCCGGAGGCGCTATGAGAATGATGCGCTGCGCATGAGC----- 950
Db 361 GAATTAATGAAGAAAGTTCCGGAGGCGCTATGAGAATGATGCGCTGCGCATGAGC----- 420
951 TCTCAGCTGATCCCTGGCCCTGAACCTGGCTGTAGTCTTTTCCAGCTTCCATCCAGC 1010
Db 421 TCTCAGCTGATCCCTGGCCCTGAACCTGGCTGTAGTCTTTTCCAGCTTCCATCCAGC 480
1011 GCAGTCCCGCGGCTCCCGAGCGGTTACTGGGGCTGCTCCCTATAGTCTCTTTATGCGAG 1070
Db 481 GCAGTCCCGCGGCTCCCGAGCGGTTACTGGGGCTGCTCCCTATAGTCTCTTTATGCGAG 540
1071 GCTCCCGAGCAGAGATGTTGCGAGGTTTATCCCGCCCGAGGAGTGGGCGGCATCATC 1130
Db 541 GCTCCCGAGCAGAGATGTTGCGAGGTTTATCCCGCCCGAGGAGTGGGCGGCATCATC 600
1131 GGCAAGAGGGGCGAGACATCAACAGCTCTCCCGTTTCCAGCGCTCCATCAAGATT 1190
Db 601 GGCAAGAGGGGCGAGACATCAACAGCTCTCCCGTTTCCAGCGCTCCATCAAGATT 660
1191 GCACCAACCCGAAACACCTGACTCCAAAGTTTCGTATGTTATCATCTGACGCGCCAGAG 1250
Db 661 GCACCAACCCGAAACACCTGACTCCAAAGTTTCGTATGTTATCATCTGACGCGCCAGAG 720
1251 GCCCAATCAAGGCTCAGGAGAAATCTATGGCAAACTCAAGAGGAGAACTCTTTTGGT 1310
Db 721 GCCCAATCAAGGCTCAGGAGAAATCTATGGCAAACTCAAGAGGAGAACTCTTTTGGT 780

1311 CCCAAGGAGGAAGTGAAGCTGGAGAGCCACATACGTGTCCAGCATCAGCAGCTGCGCGG 1370
Db 781 CCCAAGGAGGAAGTGAAGCTGGAGAGCCACATACGTGTCCAGCATCAGCAGCTGCGCGG 840
1371 GTCATTGGCAAAAGGTGCAAAACGGTGAACGAGTTCAGAAATTTGACGGCAGCTGAGGTG 1430
Db 841 GTCATTGGCAAAAGGTGCAAAACGGTGAACGAGTTCAGAAATTTGACGGCAGCTGAGGTG 900
1431 GTAGTACCAAGAGACCAAGCCCTGATGAGAACGACGACGAGTTCATCTGAAAAATCATCGGA 1490
Db 901 GTAGTACCAAGAGACCAAGCCCTGATGAGAACGACGACGAGTTCATCTGAAAAATCATCGGA 960
1491 CATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCTCTGCCCCAGGTTAAG 1550
Db 961 CATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCTCTGCCCCAGGTTAAG 1020
1551 CAGCAGCATCAGAAAGGACAGAGTAACACGAGGCCAGGACGAGGAAGTGAACGAGCCCT 1610
Db 1021 CAGCAGCATCAGAAAGGACAGAGTAACACGAGGCCAGGACGAGGAAGTGAACGAGCCCT 1080
1611 CCTGTCTCTTNGAGTCCAGGACAAACACGGGCGAGAAATCGAGAGTGTCTCTCCCGGC 1670
Db 1081 CCTGTCTCTTNGAGTCCAGGACAAACACGGGCGAGAAATCGAGAGTGTCTCTCTCCCGGC 1140
1671 AGSCCTGGAATCAGTGGGAATCCGGGACACNTGGCCGGCTGTAGATCAGGTTTGCCC 1730
Db 1141 AGSCCTGGAATCAGTGGGAATCCGGGACACCTGGCCGGCTGTAGATCAGGTTTGCCC 1200
1731 ACTTGATTGAGAAAGATGTTCCAGTGAGGAACCTGATCTNTCAGCCCCCAACACCCACC 1790
Db 1201 ACTTGATTGAGAAAGATGTTCCAGTGAGGAACCTGATCTNTCAGCCCCCAACACCCACC 1260
1791 CAATTGGCCCAACACTGTTNTGCCCTCGGGGTGTAGAAATTTNAGCGCAAGGCACTTTT 1850
Db 1261 CAATTGGCCCAACACTGTTNTGCCCTCGGGGTGTAGAAATTTNAGCGCAAGGCACTTTT 1320
1851 AAACGTGGATTGTTTAAAGAGCTCTCCAGGCCCAACGAGGAGGTTGATCACACCTCAG 1910
Db 1321 AAACGTGGATTGTTTAAAGAGCTCTCCAGGCCCAACGAGGAGGTTGATCACACCTCAG 1380
1911 TGGGAAGAAAAATAAAATTTCTTTCAGGTTTAAAAA 1946
Db 1381 TGGGAAGAAAAATAAAATTTCTTTCAGGTTTAAAAA 1416
RESULT 11
ARI60244
LOCUS 2224 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 1 from patent US 6255055.
ACCESSION ARI60244
VERSION ARI60244.1 GI:16223887
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2224)
AUTHORS Ross,J.
TITLE c-myc coding region determinant-binding protein (CRD-BP) and its
nucleic acid sequence
JOURNAL Patent: US 6255055-A 1 03-JUL-2001;
FEATURES Location/Qualifiers
source
1. .2224
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 66.9%; Score 1301; DB 6; Length 2224;
Best Local Similarity 83.5%; Pred. No. 1.9e-308;
Matches 1556; Conservative 0; Mismatches 275; Indels 33; Gaps 6;
108 AGGAGCGGAAATTCAAATCCGAATATTCACCCAGCTCCGATGGGAAGTACTGGAC 167

PI TETSUJI OTSUKI,HISASHI KOGA
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10,
PC, C12P21/02,C12Q1/68/C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT CDS (112)..(1068).
source 1..2780
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 71.4%; Score 1389; DB 6; Length 2780;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1405; Conservative 0; Mismatches 5; Indels 6; Gaps 1;
537 ACAAACAGACCCAGTCCAGATAGACGTGATAGGAAGGAGAACGCGAGTGCAGCTGAA 596
Db 1 ACAAACAGACCCAGTCCAGATAGACGTGATAGGAAGGAGAACGCGAGTGCAGCTGAA 60
597 AAAGCCATCAGTGTGCATCCACCCCTGAGGGCTGCTCCGCTTGTAAATGATCTTG 656
Qy
61 AAAGCCATCAGTGTGCATCCACCCCTGAGGGCTGCTCCGCTTGTAAATGATCTTG 120
Db
657 GAGATTATGATAAAGAGGCTTAAGACACCAAAACGGCTGACGAGGTTCCCTCGAAGATC 716
Qy
121 GAGATTATGATAAAGAGGCTTAAGACACCAAAACGGCTGACGAGGTTCCCTCGAAGATC 180
Db
717 CTGCCCCATAATACTTTGTAGGGCGTCTCATTTGCAAGAGGAGCGAACTGGAAG 776
Qy
181 CTGCCCCATAATACTTTGTAGGGCGTCTCATTTGCAAGAGGAGCGAACTGGAAG 240
Db
777 GTAGAGCAAGATACCGAGACAAAATCACCATCTCCTCGTTGCAAGACCTTACCTTTAC 836
Qy
241 GTAGAGCAAGATACCGAGACAAAATCACCATCTCCTCGTTGCAAGACCTTACCTTTAC 300
Db
837 AACCTGAGAGGACCATCAGTGTGAAGGGGCCCATCGAGAAATTTGTGAGGGCCGAGCAG 896
Qy
301 AACCTGAGAGGACCATCAGTGTGAAGGGGCCCATCGAGAAATTTGTGAGGGCCGAGCAG 360
Db
897 GAAATAATGAAGAAAGTTCCGGAGGCCATGAGAAATGATGGCTGCCATGAGC----- 950
Qy
361 GAAATAATGAAGAAAGTTCCGGAGGCCATGAGAAATGATGGCTGCCATGAGCCTGCAG 420
Db
951 TCTCACCTGATCCCTGGGCTGAACTGGCTGCTAGGCTCTTTCCGAGCTTCATCCAGC 1010
Qy
421 TCTCACCTGATCCCTGGGCTGAACTGGCTGCTAGGCTCTTTCCGAGCTTCATCCAGC 480
Db
1011 GCAGTCCCGCGCTCCAGAGCGGTTACTGGGGCTGCTCCTATAGCTCCTTTATGCGAG 1070
Qy
481 GCAGTCCCGCGCTCCAGAGCGGTTACTGGGGCTGCTCCTATAGCTCCTTTATGCGAG 540
Db
1071 GCTCCCGAGCAGGAGTGTGAGTGTATTATCCCGCCAGGCGAGTGGGCGCCATATC 1130
Qy
541 GCTCCCGAGCAGGAGTGTGAGTGTATTATCCCGCCAGGCGAGTGGGCGCCATATC 600
Db
1131 GGCAGAGAGGGCAGCAGATCAACAGCTCTCCCGTTTGGCAGCGCTCCATCAAGATT 1190
Qy
601 GGCAGAGAGGGCAGCAGATCAACAGCTCTCCCGTTTGGCAGCGCTCCATCAAGATT 660
Db
1191 GCACCACCCGAAACACCTGACTCCAAAGTTGATGTTATCATCACTGGACCGCCAGAG 1250
Qy
661 GCACCACCCGAAACACCTGACTCCAAAGTTGATGTTATCATCACTGGACCGCCAGAG 720
Db
1251 GCCCAATTCAAGGCTCAGGGAAGATCTATGGCAAACTCAAGAGGAGAACTTCTTTGGT 1310
Qy
721 GCCCAATTCAAGGCTCAGGGAAGATCTATGGCAAACTCAAGAGGAGAACTTCTTTGGT 780
Db
1311 CCCAAGGAGGAGTCAAGCTGGAGGCCACATACGTGTGCCAGATCAGCAGCTGCCGG 1370
Qy

Db 781 CCCAAGGAGGAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCTGCCGG 840
Qy 1371 GTCATTGGCAAGGTGGAATAACCGGTGAACGAGTTGAGAAATTTGACCGGAGCTGAGGTG 1430
Db 841 GTCATTGGCAAGGTGGAATAACCGGTGAACGAGTTGAGAAATTTGACCGGAGCTGAGGTG 900
Qy 1431 GTAGTACCAAGAGACGAGACCCCTGATGAGAACGACGAGTTCATCGTGAATCATCGGA 1490
Db 901 GTAGTACCAAGAGACGAGACCCCTGATGAGAACGACGAGTTCATCGTGAATCATCGGA 960
Qy 1491 CATTTCTATGCGAGTTCAGATGGCTCAACGGAAGATCCGAGACATCCTGGCCCGAGTTAAG 1550
Db 961 CATTTCTATGCGAGTTCAGATGGCTCAACGGAAGATCCGAGACATCCTGGCCCGAGTTAAG 1020
Qy 1551 CAGCAGCATCAGAAAGGAGACAGATAACGAGCCCGAGGACGAGGAAGTACCGACCCCT 1610
Db 1021 CAGCAGCATCAGAAAGGAGACAGATAACGAGCCCGAGGACGAGGAAGTACCGACCCCT 1080
Qy 1611 CCCTGTCCCTTTNGAGTCCAGGACACACGAGGCGAGAAATCGAGAGTGTCTCTCCCGCGC 1670
Db 1081 CCCTGTCCCTTTNGAGTCCAGGACACACGAGGCGAGAAATCGAGAGTGTCTCTCCCGCGC 1140
Qy 1671 AGGCTGAGAAATGAGTGGGAATCCGGGACACNTGGGCGGGCTGTAGATCAGGTTTGCC 1730
Db 1141 AGGCTGAGAAATGAGTGGGAATCCGGGACACCTGGGCGGGCTGTAGATCAGGTTTGCC 1200
Qy 1731 ACTTGATTGAGAAAGATGTTCCAGTGAGGAACCCCTGATCTWTACGCCCAACACCCACC 1790
Db 1201 ACTTGATTGAGAAAGATGTTCCAGTGAGGAACCCCTGATCTCTCAGCCCCAAACACCCACC 1260
Qy 1791 CAATTGGCCCAACACTGTTGCGGCTCGGGGTGTCAGAAATTTAGCCGACGAGCACTTTT 1850
Db 1261 CAATTGGCCCAACACTGTTGCGGCTCGGGGTGTCAGAAATTTCTAGCGCAGAGCACTTTT 1320
Qy 1851 AAACGTGATTGTTTAAAGAAAGCTCTCCAGGCCCCCACCAGAGGGTGGATCACACTCAG 1910
Db 1321 AAAGTGAGATTGTTTAAAGAAAGCTCTCCAGGCCCCCACCAGAGGGTGGATCACACTCAG 1380
Qy 1911 TGGGAAGAAAAATAAAATTTCTTCAAGTTTTTAAAA 1946
Db 1381 TGGGAAGAAAAATAAAATTTCTTCAAGTTTTTAAAA 1416
RESULT 10
AK074915 2780 bp mRNA linear PRI 03-SEP-2002
LOCUS Homo sapiens cDNA FLJ90434 fis, clone NT2RP3000789, highly similar
DEFINITION to Homo sapiens IGF-II mRNA-binding protein 1 (IMP-1) mRNA.
ACCESSION AK074915
VERSION AK074915.1 GI:22760672
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Isogai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T.,
Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S.,
Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y.,
Kojima,S., Nagahara,K., Masuho,Y., Ono,T., Okano,K., Yoshikawa,Y.,
Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and
Ninomiya,K.
NEDO human cDNA sequencing project
TITLE
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2780)
AUTHORS Isogai,T. and Otsuki,T.
DIRECT SUBMISSION
TITLE
JOURNAL Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomicshri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:

ORIGIN		/db_xref="taxon:9606"	
Query Match		85.8%; Score 1669.2; DB 6; Length 1708;	
Best Local Similarity		99.8%; Pred. No. 0;	
Matches 1676; Conservative		0; Mismatches 3; Indels 0; Gaps 0;	
QY	268	CCAGGCAAGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCCTGAAGGTCT	327
Db	30	CCCGGGAGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCCTGAAGGTCT	89
QY	328	CTTACATCCCGATGACGATAGCACAGGACCTGAGATGGCGCGGAGGGGGCTTG	387
Db	90	CTTACATCCCGATGACGATAGCACAGGACCTGAGATGGCGCGGAGGGGGCTTG	149
QY	388	GCTCTCGGGTCAGCCCGCCGAGGGCTCACCTGTGGCAGCGGGGGCCCGACCAAGCAGC	447
Db	150	GCTCTCGGGTCAGCCCGCCGAGGGCTCACCTGTGGCAGCGGGGGCCCGACCAAGCAGC	209
QY	448	AGCAAGTGGACATCCCGCTTCGGCTCTGTGTGCCACCCAGTATGTGGGTGCCATTATG	507
Db	210	AGCAAGTGGACATCCCGCTTCGGCTCTGTGTGCCACCCAGTATGTGGGTGCCATTATG	269
QY	508	GCAGAGGGGGCACCATCCGCAACATCACAAACAGACCCAGTCCAGATAGAGCTGC	567
Db	270	GCAGAGGGGGCACCATCCGCAACATCACAAACAGACCCAGTCCAGATAGAGCTGC	329
QY	568	ATAGGAAGAGAAACGAGGTGCAAGCTGAAAGGCAATCAGTGTGCATCCACCCCTGAGG	627
Db	330	ATAGGAAGAGAAACGAGGTGCAAGCTGAAAGGCAATCAGTGTGCATCCACCCCTGAGG	389
QY	628	GCTGCTCTCCGCTTGAAGATGATCTGGAGATATATGCAATAAGAGGCTTAAGACCA	687
Db	390	GCTGCTCTCCGCTTGAAGATGATCTGGAGATATATGCAATAAGAGGCTTAAGACCA	449
QY	688	AAACGGCTGACAGGTTCCCTGAAGATCTGTGCCCATATATCTTTGTAGGGCGTCTCA	747
Db	450	AAACGGCTGACAGGTTCCCTGAAGATCTGTGCCCATATATCTTTGTAGGGCGTCTCA	509
QY	748	TTGGCAAGAGGACGGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAAAATCACCA	807
Db	510	TTGGCAAGAGGACGGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAAAATCACCA	569
QY	808	TCCTCTGTTGGAAGACTTACCTTTACAACTGTAGAGGACATCATCTGTGAAGGGGG	867
Db	570	TCCTCTGTTGGAAGACTTACCTTTACAACTGTAGAGGACATCATCTGTGAAGGGGG	629
QY	868	CCATCGAAGATTGTTCAGGGCCGAGCAGGAAATATGAAGAAAGTTTCGGGAGGCCCTATG	927
Db	630	CCATCGAAGATTGTTCAGGGCCGAGCAGGAAATATGAAGAAAGTTTCGGGAGGCCCTATG	689
QY	928	AGAATGATGTGGTGCATGAGCTCTCACCTGATCCCTGGCCTGAACCTGGCTGTAG	987
Db	690	AGAATGATGTGGTGCATGAGCTCTCACCTGATCCCTGGCCTGAACCTGGCTGTAG	749
QY	988	GTCTTTTCCAGCTTCTACAGCGCAGTCCCGCGCCTCCAGCAGGCTTACTGGGGCTG	1047
Db	750	GTCTTTTCCAGCTTCTACAGCGCAGTCCCGCGCCTCCAGCAGGCTTACTGGGGCTG	809
QY	1048	CTCCCTATAGCTCTTTATGAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTATCCCGG	1107
Db	810	CTCCCTATAGCTCTTTATGAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTATCCCGG	869
QY	1108	CCAGGAGTGGGCGCCATCATCGGCAAGAGGGGAGCAGCATCAACAGCTCTCCCGGT	1167
Db	870	CCAGGAGTGGGCGCCATCATCGGCAAGAGGGGAGCAGCATCAACAGCTCTCCCGGT	929
QY	1168	TTGCGAGGCTTCATCAAGATTGACACCCGAAACACCTGACTCCAAAGTTGTTATGG	1227
Db	930	TTGCGAGGCTTCATCAAGATTGACACCCGAAACACCTGACTCCAAAGTTGTTATGG	989
QY	1228	TTATCATCACTGGACCGCAGGCCCAATTCAGGCTCAGGGAAGATCTATGGCAAC	1287
Db			

990	TTATCATCACTGGACCGCCAGAGGGCCCAATTCAGGCTCAGGGAAGAAATCTATGGCAAC	1049
1288	TCAAGGAGGAGAACTTTCTTTTGGTCCCAAGGAGAGTGAAGCTGGAGACCCACATACGTG	1347
1050	TCAAGGAGGAGAACTTTCTTTTGGTCCCAAGGAGAGTGAAGCTGGAGACCCACATACGTG	1109
1348	TGCCAGCATCAGCAGCTGCGCGGTCTATTTGGCAAGGTGGAAACCGTGAACGAGTTGC	1407
1110	TGCCAGCATCAGCAGCTGCGCGGTCTATTTGGCAAGGTGGAAACCGTGAACGAGTTGC	1169
1408	AGAAATTTGACGCGCAGCTGAGGTGGTAGTACCAAGAGACAGACCCCTGATGAGAACGACC	1467
1170	AGAAATTTGACGCGCAGCTGAGGTGGTAGTACCAAGAGACAGACCCCTGATGAGAACGACC	1229
1468	AGTCACTCGTGAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCC	1527
1230	AGTCACTCGTGAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCC	1289
1528	GAGACATCTCGGCCAGGTTAAGCAGCAGCATCAGAGGACAGATTAACGAGGCCAGG	1587
1290	GAGACATCTCGGCCAGGTTAAGCAGCAGCATCAGAGGACAGATTAACGAGGCCAGG	1349
1588	CACGAGGAGAGTGACCCAGCCCTCCCTGTCCCTTNGAGTCCAGGACAAACACGGGACAGAA	1647
1350	CACGAGGAGAGTGACCCAGCCCTCCCTGTCCCTTNGAGTCCAGGACAAACACGGGACAGAA	1409
1648	ATCGAGAGTGTCTCTCTCCCGGAGCCCTGAGAAATGAGTGGGAATCCGGGACACNTGGGC	1707
1410	ATCGAGAGTGTCTCTCTCCCGGAGCCCTGAGAAATGAGTGGGAATCCGGGACACNTGGGC	1469
1708	CGGGCTGTAGATCAGGTTTGGCCACTTGTGAAAGATGTTCCAGTCAGGAAACCTGGA	1767
1470	CGGGCTGTAGATCAGGTTTGGCCACTTGTGAAAGATGTTCCAGTCAGGAAACCTGGA	1529
1768	TCNTCAGCCCAACACCCACCAATTCGCCCAACACTGTNTGCCCTCGGGGTGTGAG	1827
1530	TCNTCAGCCCAACACCCACCAATTCGCCCAACACTGTNTGCCCTCGGGGTGTGAG	1589
1828	AAATTNTAGCGAAGCATTCTTTAAACGTGGATGTTTAAAGAGCTCTCCAGGCCCCAC	1887
1590	AAATTNTAGCGAAGCATTCTTTAAACGTGGATGTTTAAAGAGCTCTCCAGGCCCCAC	1649
1888	CAAGAGGTTGATCAGACCTCAGTGGGAGAGAAATAAATTTCTCAGTTTAAAA	1946
1650	CAAGAGGTTGATCAGACCTCAGTGGGAGAGAAATAAATTTCTCAGTTTAAAA	1708

BD127811	2780 bp	DNA	linear	PAT 18-SEP-2002
Primer for synthesizing full-length cDNA and use thereof.				
BD127811				
BD127811.1	GI:23222756			
JP 2002017375-A/3242				
Homo sapiens (human)				
Homo sapiens				
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y., Wakanatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and Koga, H.				
1 (bases 1 to 2780)				
Primer for synthesizing full-length cDNA and use thereof				
Patent: JP 2002017375-A 3242 22-JAN-2002;				
HELIX RESEARCH INSTITUTE				
OS Homo sapiens (human)				
PN JP 2002017375-A/3242				
PD 22-JAN-2002				
PF 07-JUL-2000 JP 2000253172				
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO ISHII,				
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI SHINICHI KOJIMA,				

QY	388	GCTCTCGGGGT	CAGCCCCCGCCAGGGCT	CACCTGTGGCAGCGGGGGCCCCAGCAAGCAGC	447
Db	150	GCTCTCGGGGT	CAGCCCCCGCCAGGGCT	CACCTGTGGCAGCGGGGGCCCCAGCAAGCAGC	209
QY	448	AGCAAGTGGACAT	CCCCCTTCGGCT	CCTGGTGCACCCACCCAGTATGTGGGTGCCATTATTG	507
Db	210	AGCAAGTGGACAT	CCCCCTTCGGCT	CCTGGTGCACCCACCCAGTATGTGGGTGCCATTATTG	269
QY	508	GCAAGGAGGGGCC	CACCATCCGCAACATC	CAAAAAAGACCCAGTCCAGATAGACGTGC	567
Db	270	GCAAGGAGGGGCC	CACCATCCGCAACATC	CAAAAAAGACCCAGTCCAGATAGACGTGC	329
QY	568	ATAGGAAGGAGAAC	CGCAGGTGCAGCT	GAAAAAGCCATCAGTGTGCATCCACCCCTGAGG	627
Db	330	ATAGGAAGGAGAAC	CGCAGGTGCAGCT	GAAAAAGCCATCAGTGTGCATCCACCCCTGAGG	389
QY	628	GCTGCTCCTCGCT	TGTAAGATGATCT	TGAGATTATGATAAAGAGGCTTAAGGACACCA	687
Db	390	GCTGCTCCTCGCT	TGTAAGATGATCT	TGAGATTATGATAAAGAGGCTTAAGGACACCA	449
QY	688	AAACGGCTGACGAG	TTCCCTGAAGATCCT	TGGCCCATATAACTTTGTAGGGCGTCTCA	747
Db	450	AAACGGCTGACGAG	TTCCCTGAAGATCCT	TGGCCCATATAACTTTGTAGGGCGTCTCA	509
QY	748	TTGCAAGGAGGAC	CGGAACCTGAAGAGGT	TAGACAGATACCGGACAAAAAATCACCA	807
Db	510	TTGCAAGGAGGAC	CGGAACCTGAAGAGGT	TAGACAGATACCGGACAAAAAATCACCA	569
QY	808	TCTCCTCGTTGCA	AGACTTTACCTTTAC	CAACCTTGAGAGGACATCAGTGTGAAGGGG	867
Db	570	TCTCCTCGTTGCA	AGACTTTACCTTTAC	CAACCTTGAGAGGACATCAGTGTGAAGGGG	629
QY	868	CCATCGAGAAATGT	TTCAGGCGCAGCAGAA	TAATGAAGAAAGTTTCGGGAGGCGCTATG	927
Db	630	CCATCGAGAAATGT	TTCAGGCGCAGCAGAA	TAATGAAGAAAGTTTCGGGAGGCGCTATG	689
QY	928	AGAAATGATGGCT	GCACATGAGCTCT	CACCTGATCCCTGGCCCTGAACCTGGCTGCTGTAG	987
Db	690	AGAAATGATGGCT	GCACATGAGCTCT	CACCTGATCCCTGGCCCTGAACCTGGCTGCTGTAG	749
QY	988	GTCTTTTCCAGCT	TTCAATCCAGCGAGT	CCGCGCGCTCCAGCAGAGGTTACTTGGGGCTG	1047
Db	750	GTCTTTTCCAGCT	TTCAATCCAGCGAGT	CCGCGCGCTCCAGCAGAGGTTACTTGGGGCTG	809
QY	1048	CTCCCTATAGCT	CTCTTTATCAGGCT	CCCGAGAGGAGATGGTGCAGTGTTTATCCCGG	1107
Db	810	CTCCCTATAGCT	CTCTTTATCAGGCT	CCCGAGAGGAGATGGTGCAGTGTTTATCCCGG	869
QY	1108	CCCAGGAGTGGG	CGCCATCATCGCAAGAG	GGGGCAGCATCAAAACAGCTCTCCCGGT	1167
Db	870	CCCAGGAGTGGG	CGCCATCATCGCAAGAG	GGGGCAGCATCAAAACAGCTCTCCCGGT	929
QY	1168	TTGCCAGGCGCT	CCATCAGANTTGCAC	CAACCGGAGTCCAAAGTTCGTATGG	1227
Db	930	TTGCCAGGCGCT	CCATCAGANTTGCAC	CAACCGGAGTCCAAAGTTCGTATGG	989
QY	1228	TTATCATCACTGG	ACCGCCAGAGCCCAAT	TCAAGGCTCAGGGGAAGATCTATGGCAAC	1287
Db	990	TTATCATCACTGG	ACCGCCAGAGCCCAAT	TCAAGGCTCAGGGGAAGATCTATGGCAAC	1049
QY	1288	TCAAGGAGGAGAA	CTTCTTTTGGTCCC	CAAGGAGGAAGTGAAGCTGGAGACCCACATACGTG	1347
Db	1050	TCAAGGAGGAGAA	CTTCTTTTGGTCCC	CAAGGAGGAAGTGAAGCTGGAGACCCACATACGTG	1109
QY	1348	TGCCAGCATCAG	CAGCTGGCGGGTCA	TGGCAAAAGTGGAAAAACCGTGAAACGAGTTGC	1407
Db	1110	TGCCAGCATCAG	CAGCTGGCGGGTCA	TGGCAAAAGTGGAAAAACCGTGAAACGAGTTGC	1169
QY	1408	AGAAATTCACGG	CAGCTGAGTGTGTAG	TACCAAGAGACGAGCCCTGATGAGAACGACC	1467
Db	1170	AGAAATTCACGG	CAGCTGAGTGTGTAG	TACCAAGAGACGAGCCCTGATGAGAACGACC	1229
QY	1468	AGGTCATCGT	GAATAATCATCGACAT	TTCTATGCCAGTTCAGATGGCTCAACGGGAATGCC	1527

Db	1230	AGGTCATCGTGAAAAAATCATCGGACATTTCTATGCCAGTCCAGATGGCTCAACGGAAGATCC	1289
Qy	1528	GAGACATCTCTGGCCAGGTAAAGCAGCAGCATCAGAAGGGGACAGAGTAACACGAGCCAGG	1587
Db	1290	GAGACATCTCTGGCCAGGTAAAGCAGCAGCATCAGAAGGGGACAGAGTAACACGAGCCAGG	1349
Qy	1588	CACGGAGGAGTAGCACACAGCCCCTCTCTGTCCTCTTNGAGTCCAGGACAAACAACGGGCAGAA	1647
Db	1350	CACGGAGGAGTAGCACACAGCCCCTCTCTGTCCTCTTNGAGTCCAGGACAAACAACGGGCAGAA	1409
Qy	1648	ATCAGAGGTGTCTCTCCCGGACAGCCCTGAGAAATCAGTGGGAATCCCGGACACANTGGGC	1707
Db	1410	ATCAGAGGTGTCTCTCCCGGACAGCCCTGAGAAATCAGTGGGAATCCCGGACACANTGGGC	1469
Qy	1708	CGGGCTGTAGATCAGAGTTTGCCCACTTGATTGAGAAAGATGTTCCAGTGAAGAACCTTGA	1767
Db	1470	CGGGCTGTAGATCAGAGTTTGCCCACTTGATTGAGAAAGATGTTCCAGTGAAGAACCTTGA	1529
Qy	1768	TCTNTCAGCCCCAAAACACCACCCCAATTCGGCCCAACACTGTGTGTCCTCGGGGTGTGAG	1827
Db	1530	TCTNTCAGCCCCAAAACACCACCCCAATTCGGCCCAACACTGTGTGTCCTCGGGGTGTGAG	1589
Qy	1828	AAATTNTAGCGCAAGCACTTTTAAACGTGGATTGTTTAAAGAGCTCTCCAGGCCCCAC	1887
Db	1590	AAATTNTAGCGCAAGCACTTTTAAACGTGGATTGTTTAAAGAGCTCTCCAGGCCCCAC	1649
Qy	1888	CAAGAGGGTGGATCACAACCTCAGTGGGAAGAAAATAAAATTTCTTCAGGTTTTAAAA	1946
Db	1650	CAAGAGGGTGGATCACAACCTCAGTGGGAAGAAAATAAAATTTCTTCAGGTTTTAAAA	1708
RESULT 8			
BD209924		1708 bp DNA linear PAT 17-JUL-2003	
LOCUS		Isolated nucleic acid molecules encoding cancer-associated	
DEFINITION		antigens, these antigens and method of using the same.	
ACCESSION		BD209924	
VERSION		BD209924.1 GI:33019694	
KEYWORDS		JP 2002512049-A/3.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS		Chen,Y.T., Gure,A., Tsang,S., Stockert,E., Jager,E., Knuth,A. and Old,L.J.	
TITLE		Isolated nucleic acid molecules encoding cancer-associated	
JOURNAL		antigens, these antigens and method of using the same	
COMMENT		Patent: JP 2002512049-A 3 23-APR-2002; LUDWIG INSTITUTE FOR CANCER RESEARCH	
		OS Homo sapiens (human)	
		PN JP 2002512049-A/3	
		PD 23-APR-2002	
		PF 16-MAR-1999 JP 2000545030	
		PR 17-APR-1998 US 09/061709	
		PI YAO TSENG CHEN,ALI GURE,SOLAM TSANG,ELISABETH STOCKERT,ELKE	
		PI JAGER,	
		PI ALEXANDER KNUTH,LLOYD J OLD	
		PC C12N15/09,A61K35/12,A61K39/00,A61P35/00,C07K16/32,	
		PC C12N1/15,	
		PC C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/53,G01N33/ PC	
		574,	
		PC C12N15/00,C12N5/00	
		CC n=not determined(1384,1464,1533,1571,1595)	
		FH key Location/Qualifiers	
		FT source 1..1708	
		FT Location/Qualifiers /organism='Homo sapiens (human)'	
		1..1708 /organism='Homo sapiens'	
		/mol_type='genomic DNA'	

Db 30 CCGGGAGGACATCATGAAGCTGATGGCCACAGTTGGAGAACCATGGCCCTGAAGGTCT 89
QY 328 CTTACATPCCCGATGAGCAGATAGCA CAGGAGACCTGAGAATGGCGCGCGAGGGGGCTTTG 387
Db 90 CTTACATPCCCGATGAGCAGATAGCA CAGGAGACCTGAGAATGGCGCGCGAGGGGGCTTTG 149
QY 388 GCTCTCGGGGTGAGCCCGCGAGGGCTCACTGTGGCAGAGGGGGCCCGAGCCGACAGCAGC 447
Db 150 GCTCTCGGGGTGAGCCCGCGAGGGCTCACCTGTGGCAGAGGGGGCCCGAGCCGACAGCAGC 209
QY 448 AGCAAGTGAACATCCCGCTTCGGCTCTGCTGTGGCCACCCAGTATGTGGGTGCCATTATTG 507
Db 210 AGCAAGTGAACATCCCGCTTCGGCTCTGCTGTGGCCACCCAGTATGTGGGTGCCATTATTG 269
QY 508 GCAAGAGGGGGGACCATCGCAACATCA CAAACAGACCCAGTCCAAAGATAGACGTGC 567
Db 270 GCAAGAGGGGGGACCATCGCAACATCA CAAACAGACCCAGTCCAAAGATAGACGTGC 329
QY 568 ATAGGAGGAGGAGCGAGGTGCAGCTGAAAAGCCATCAGTGTGCATCTCCACCCCTGAGG 627
Db 330 ATAGGAGGAGGAGCGAGGTGCAGCTGAAAAGCCATCAGTGTGCATCTCCACCCCTGAGG 389
QY 628 GCTGCTCTCCGCTTGTAAAGATGATCTTGGAGATTATGCATAAAGAGGCTTAAGCACCA 687
Db 390 GCTGCTCTCCGCTTGTAAAGATGATCTTGGAGATTATGCATAAAGAGGCTTAAGCACCA 449
QY 688 AAGCGGTGACGAGGTTCCTCGAAGTCTCGGCCATAATAAATTTGTAGGGCGTCTCA 747
Db 450 AAGCGGTGACGAGGTTCCTCGAAGTCTCGGCCATAATAAATTTGTAGGGCGTCTCA 509
QY 748 TTGCGAAGGAGGACGAACTCGAAGAGTGTAGGCAAGATACCGAGACAAAATCACCA 807
Db 510 TTGCGAAGGAGGACGAACTCGAAGAGTGTAGGCAAGATACCGAGACAAAATCACCA 569
QY 808 TCTCCTCGTTCGAAGACCTTACCCCTTTACAAACCTGTAGAGGACCATCACCTGTGAAGGGG 867
Db 570 TCTCCTCGTTCGAAGACCTTACCCCTTTACAAACCTGTAGAGGACCATCACCTGTGAAGGGG 629
QY 868 CCAATGAGAAATTTGTCAGGGCGGAGCAGGAAATAATGAAGAAAGTTTCGGAGGCTTATG 927
Db 630 CCAATGAGAAATTTGTCAGGGCGGAGCAGGAAATAATGAAGAAAGTTTCGGAGGCTTATG 689
QY 928 AGAATGATGCGTGCATGAGCTCTCACCTGATCCCTGGCTGAACTGGCTGCTGTAG 987
Db 690 AGAATGATGCGTGCATGAGCTCTCACCTGATCCCTGGCTGAACTGGCTGCTGTAG 749
QY 988 GTCTTTTCCAGCTTCATCCAGGCGAGTCCCGCGGCTCCAGCAGCGTTACTTGGGGCTG 1047
Db 750 GTCTTTTCCAGCTTCATCCAGGCGAGTCCCGCGGCTCCAGCAGCGTTACTTGGGGCTG 809
QY 1048 CTCCTATAGCTCTTTATGAGGCTCCCGAGCAGAGATGGTGCAGGTGTTATCCCG 1107
Db 810 CTCCTATAGCTCTTTATGAGGCTCCCGAGCAGAGATGGTGCAGGTGTTATCCCG 869
QY 1108 CCCAGGCGTGGCGGCTCATCGCAAGAGGGGAGCAGATCAACAGCTCTCCCGT 1167
Db 870 CCCAGGCGTGGCGGCTCATCGCAAGAGGGGAGCAGATCAACAGCTCTCCCGT 929
QY 1168 TTGCCAGGCGCTCCATCAAGATTGCACCAACCGGAAACACCTGACTCCAAAGTTTCGTATGG 1227
Db 930 TTGCCAGGCGCTCCATCAAGATTGCACCAACCGGAAACACCTGACTCCAAAGTTTCGTATGG 989
QY 1228 TTATCATCTAGCAGCGCAGAGCCCAATTTCAAGGCTAGGGAAGATCTATGGCAAC 1287
Db 990 TTATCATCTAGCAGCGCAGAGCCCAATTTCAAGGCTAGGGAAGATCTATGGCAAC 1049
QY 1288 TCAAGAGGAGAACTCTTTTGTCCCAAGGAGGAGTGAAGCTGGAGCCCAATACGTG 1347
Db 1050 TCAAGAGGAGAACTCTTTTGTCCCAAGGAGGAGTGAAGCTGGAGCCCAATACGTG 1109
QY 1348 TGCCAGCATCAGCAGCTGGCGGGGTCACTTGGCAAGAGTGGAAAAACGGTGAACAGTTGC 1407

Db 1110 TGCCAGCATCAGCAGCTGCGCGGTCACTTGGCAAAAGGTGGAATAACGGTGAAACAGTTGC 1169
QY 1408 AGAATTTGACGCGCAGCTGAGGTGGTAGTACCAAGAGACCAAGACCCCTGATGAGAACGACC 1467
Db 1170 AGAATTTGACGCGCAGCTGAGGTGGTAGTACCAAGAGACCAAGACCCCTGATGAGAACGACC 1229
QY 1468 AGTCACTGTAAGAAATCATCGGACATTTCTATGCCAGTCAAGTGGCTCAACCGAAGATCC 1527
Db 1230 AGTCACTGTAAGAAATCATCGGACATTTCTATGCCAGTCAAGTGGCTCAACCGAAGATCC 1289
QY 1528 GACACATCTTGGCCCGAGGTAAAGCAGCAGCATCAGAGGAGCAGAGTAACACGAGCCGAGG 1587
Db 1290 GAGACATCTTGGCCCGAGGTAAAGCAGCAGCATCAGAGGAGCAGAGTAACACGAGCCGAGG 1349
QY 1588 CACGAGGAGTGAACACGAGCCCTCCCTGCTTCCCTTNGAGTCCAGGACAAACAAACGGGAGAA 1647
Db 1350 CACGAGGAGTGAACACGAGCCCTCCCTGCTTCCCTTNGAGTCCAGGACAAACAAACGGGAGAA 1409
QY 1648 ATCAGAGTGTGCTCTCCCGGAGGCGCTGAGATGAGTGGGAATCGGGACACNTGGGC 1707
Db 1410 ATCAGAGTGTGCTCTCCCGGAGGCGCTGAGATGAGTGGGAATCGGGACACNTGGGC 1469
QY 1708 CGGCTGTAGATCAGGTTCGCCACTTGTGATGAGAAAGATGTTCCAGTGAAGAACCTGA 1767
Db 1470 CGGCTGTAGATCAGGTTCGCCACTTGTGATGAGAAAGATGTTCCAGTGAAGAACCTGA 1529
QY 1768 TCTNTCAGCCCCAAAACACCCCAATTCGCCCCAACACTGTNTGCCCTCGGGGTGTGAG 1827
Db 1530 TCTNTCAGCCCCAAAACACCCCAATTCGCCCCAACACTGTNTGCCCTCGGGGTGTGAG 1589
QY 1828 AAATTTAGCGCAAGGACCTTTTAAACGTTGGATGTTTAAAGAGCTCTCCAGGCCCCAC 1887
Db 1590 AAATTTAGCGCAAGGACCTTTTAAACGTTGGATGTTTAAAGAGCTCTCCAGGCCCCAC 1649
QY 1888 CAAGAGGTGGATCACACCTCAGTGGGAAAGAAAATAAAATTTCCCTCAGGTTTAAAA 1946
Db 1650 CAAGAGGTGGATCACACCTCAGTGGGAAAGAAAATAAAATTTCCCTCAGGTTTAAAA 1708

RESULT 7

LOCUS AR343073 1708 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 5 from patent US 6576756.
ACCESSION AR343073
VERSION AR343073.1 GI:33738475
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1708)
AUTHORS Chen, Y.-T., Gure, A., Tsang, S., Stockert, E., Jager, E., Alexander, K.
and Old, L.J.
TITLE Isolated nucleic acid molecule encoding cancer associated antigen,
the antigen itself, and uses thereof
JOURNAL Patent: US 6576756-A 5 10-JUN-2003;
FEATURES
source location/Qualifiers
1. .1708 /organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 85.8%; Score 1669.2; DB 6; Length 1708;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1676; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 268 CCAGGCAAGCCATCATGAAGCTGAATGGCCACACAGTTGGAGAACCATGCCCTGAAGGTCT 327
Db 30 CCGGGGAGCCATCATGAAGCTGAATGGCCACACAGTTGGAGAACCATGCCCTGAAGGTCT 89
QY 328 CTTACATCCCGATGAGCAGATAGCACAGGACCTGAGAAATGGGCGCGAGGGGGCTTTG 387
Db 90 CTTACATCCCGATGAGCAGATAGCACAGGACCTGAGAAATGGGCGCGAGGGGGCTTTG 149

Db 600 AGCTGCTGGCTCAGTATGTTAGTAGAGAACTGTGAGCAAGTGAACACCGAGGTGAG 659

Qy 228 ACGCAGTGTGTAATGTCACTTATCCAAACGGGAGCAGACCGAGCAAGCCATCATGAAG 287

Db 660 ACGCAGTGTGTAATGTCACTTATCCAAACGGGAGCAGACCGAGCAAGCCATCATGAAG 719

Qy 288 CTGAATGGCCACCAAGTTGGAGAACATGCTCCCTGAAGGTCTCTCTATCATCCCGGATGAGCAG 347

Db 720 CTGAATGGCCACCAAGTTGGAGAACATGCTCCCTGAAGGTCTCTCTATCATCCCGGATGAGCAG 779

Qy 348 ATAGCAGAGGACCTGGAATGGCGCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGC 407

Db 780 ATAGCAGAGGACCTGGAATGGCGCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGC 839

Qy 408 CAGGGCTCACCTGTGGCAGCGGGGCCCCAGCAGCAGCAGCAGTGGACATCCCTTT 467

Db 840 CAGGGCTCACCTGTGGCAGCGGGGCCCCAGCAGCAGCAGCAGTGGACATCCCTTT 899

Qy 468 CGGCTCTGGTGCCACCCAGTATGTGGTGCCATTATTGGCAAGGAGGGGCCACCATC 527

Db 900 CGGCTCTGGTGCCACCCAGTATGTGGTGCCATTATTGGCAAGGAGGGGCCACCATC 959

Qy 528 CGCAACATCAAAACAGCAGCCAGTCCAAAGATAGACGTGCTATGAGGAGGAGCAGCAGT 587

Db 960 CGCAACATCAAAACAGCAGCCAGTCCAAAGATAGACGTGCTATGAGGAGGAGCAGCAGT 1019

Qy 588 GCAGCTGAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGTGCTCTCCCTCGCTTTGAAG 647

Db 1020 GCAGCTGAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGTGCTCTCCCTCGCTTTGAAG 1079

Qy 648 ATGATCTTGGAGATTATGCAATAAGAGGCTAAGGACACCAAAACGGCTGACGAGTTCC 707

Db 1080 ATGATCTTGGAGATTATGCAATAAGAGGCTAAGGACACCAAAACGGCTGACGAGTTCC 1139

Qy 708 CTGAAGATCTGGGCCATAAATCTTGTAGGGCTCTCATGGCAAGGAGGAGCGAAC 767

Db 1140 CTGAAGATCTGGGCCATAAATCTTGTAGGGCTCTCATGGCAAGGAGGAGCGAAC 1199

Qy 768 CTGAAGAGGTAGAGCAAGATACCGAGACAAATCACCATCTCTCTGCTGCAAGACCTT 827

Db 1200 CTGAAGAGGTAGAGCAAGATACCGAGACAAATCACCATCTCTCTGCTGCAAGACCTT 1259

Qy 828 ACCCTTTACACCTGAGAGACCATCATCTGTGAAGGGGGCCATCGAGAAATTTGTCAGG 887

Db 1260 ACCCTTTACACCTGAGAGACCATCATCTGTGAAGGGGGCCATCGAGAAATTTGTCAGG 1319

Qy 888 GCGAGAGGAAATTAATGAAGAAAGTTCGGGAGGCTATGAGAAATGATGTGCTGCCATG 947

Db 1320 GCGAGAGGAAATTAATGAAGAAAGTTCGGGAGGCTATGAGAAATGATGTGCTGCCATG 1379

Qy 948 AGC-----TCTCACCCTGATCCCTGGCCTGAACTGGCTGTGTAGGTCTTTTCCAGCT 1001

Db 1380 AGCCTGCAGTCTCACCCTGACCCCTGGCCTGAACTGGCTGTGTAGGTCTTTTCCAGCT 1439

Qy 1002 TCATCCAGGAGTCCCGCGCTCCAGCAGCTTACTGGGCTGCTCCCTATAGCTCC 1061

Db 1440 TCATCCAGGAGTCCCGCGCTCCAGCAGCTTACTGGGCTGCTCCCTATAGCTCC 1499

Qy 1062 TTTATGAGGCTCCCGAGCAGAGATGTTGAGGTGTTTATCCCGCCCGAGGAGTGGG 1121

Db 1500 TTTATGAGGCTCCCGAGCAGAGATGTTGAGGTGTTTATCCCGCCCGAGGAGTGGG 1559

Qy 1122 GCCATCATCGGCAAGAGGGGCGAGCAGCATCAACAGCTCTCCCGTTTCCAGCGCTCC 1181

Db 1560 GCCATCATCGGCAAGAGGGGCGAGCAGCATCAACAGCTCTCCCGTTTCCAGCGCTCC 1619

Qy 1182 ATCAGATGACGACCCGAGACAGCTGACTCAAGTTCGTATGTTATCATCTATGGA 1241

Db 1620 ATCAGATGACGACCCGAGACAGCTGACTCAAGTTCGTATGTTATCATCTATGGA 1679

Qy 1242 CCGCCAGAGGCCCAATCAAGGCTCAGGGAAGAAATCTATGGCAAACTCAAGGAGGAGAAC 1301

Db 1680 CCGCCAGAGGCCCAATCAAGGCTCAGGGAAGAAATCTATGGCAAACTCAAGGAGGAGAAC 1739

Qy 1302 TTCTTTGGTCCCAAGGAGGAGTGAAGCTGGAGACCCACATACGTTGTCGAGCATCAGCA 1361

Db 1740 TTCTTTGGTCCCAAGGAGGAGTGAAGCTGGAGACCCACATACGTTGTCGAGCATCAGCA 1799

Qy 1362 GCTGCGCGGGTCATTGGCAAAAGGTGGAACCGGTGAACGAGTTTCAGAAATTTGACGCA 1421

Db 1800 GCTGCGCGGGTCATTGGCAAAAGGTGGAACCGGTGAACGAGTTTCAGAAATTTGACGCA 1859

Qy 1422 GCTGAGGTGTAGTACCAAGAGACAGACCCCTGATGAGAACGACGACGAGTCACTGTA 1481

Db 1860 GCTGAGGTGTAGTACCAAGAGACAGACCCCTGATGAGAACGACGAGTCACTGTA 1919

Qy 1482 ATCATCGACATTTCTATGCCAGTCAAGTGTCAACGAGATCCGAGACATCTGGCC 1541

Db 1920 ATCATCGACATTTCTATGCCAGTCAAGTGTCAACGAGATCCGAGACATCTGGCC 1979

Qy 1542 CAGGTTAAGCAGCAGCATCAGAAAGGACAGAGTAACAGGGCCAGGACCGAGGAAGTGA 1601

Db 1980 CAGGTTAAGCAGCAGCATCAGAAAGGACAGAGTAACAGGGCCAGGACCGAGGAAGTGA 2039

Qy 1602 CCAGCCCTCCCTGTTCCTTNGAGTCCAGGACAAACACGGGAGAAATCGAGAGTGTGCT 1661

Db 2040 CCAGCCCTCCCTGTTCCTTNGAGTCCAGGACAAACACGGGAGAAATCGAGAGTGTGCT 2099

Qy 1662 CTCGCCGAGCGCTGAGAAATGAGTGGGAATCCGGACACANTGGCCGCGCTGTAGATCA 1721

Db 2100 CTCGCCGAGCGCTGAGAAATGAGTGGGAATCCGGACACCTGGGCGCGCTGTAGATCA 2159

Qy 1722 GGTTCGCCACCTTGATTCAGAAAGATGTTCCAGTCAGGAACCTCTGATCTNTCAG-CCCCA 1780

Db 2160 GGTTCGCCACCTTGATTCAGAAAGATGTTCCAGTCAGGAACCTCTGATCTCAGCCCCA 2219

Qy 1781 AACACCCACCAATTTGGCCCAACACTGNTGCCCTCGGGGTGTGAGAAATTTAGCGCA 1840

Db 2220 AACACCTACCAATTTGGCCCAACACTGNTGCCCTCGGGGTGTGAGAAATTTAGCGCA 2279

Qy 1841 AGGCATTTTAAACCTGATTTTAAAGAGCTCTCCAGGCCCCACCAAGAGGGTGGAT 1900

Db 2280 AGGCATTTTAAACCTGATTTTAAAGAGCTCTCCAGGCCCCACCAAGAGGGTGGAT 2339

Qy 1901 CACACCTCAGTGGGAGAAATAAATAATTTCTTCAGGTTTTT 1942

Db 2340 CACACCTCAGTGGGAGAAATAAATAATTTCTTCAGGTTTTT 2381

RESULT 6

AR171864

LOCUS AR171864 1708 bp DNA linear PAT 17-DEC-2001

DEFINITION Sequence 5 from patent US 6297364.

ACCESSION AR171864

VERSION AR171864.1 GI:17910814

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 1708)

AUTHORS Chen,Y.-T., Gure,A., Tsang,S., Stockert,E., Jager,E., Alexander,K. and Old,L.J.

TITLE Isolated nucleic acid molecule encoding cancer associated antigen, the antigen itself, and uses thereof

JOURNAL Patent: US 6297364-A 5 02-OCT-2001;

FEATURES

source Location/Qualifiers

1..1708

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Query Match 85.8%; Score 1669.2; DB 6; Length 1708;

Best Local Similarity 99.8%; Fred. No. 0;

Matches 1676; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 268 CCAGGCAAGCCATCATGAAGCTGAATGGCCACCAAGTGGAGAACCATGCCCTGAAGGCTCT 327

```
QY 768 CTGAAGAAGTAGAGCAAGATACCGAGCAAAAATCACTCTCTCTGTTGCAAGACCTT 827
Db 904 CTGAAGAAGTAGAGCAAGATACCGAGCAAAAATCACTCTCTCTGTTGCAAGACCTT 963
QY 828 ACCCTTTACAAACCTCAGAGGACCATCAGTGTGAAGGGGCCATCAGAAATGCTTCAGG 887
Db 964 ACCCTTTACAAACCTCAGAGGACCATCAGTGTGAAGGGGCCATCAGAAATGCTTCAGG 1023
QY 888 GCCGAGCAGGAAATAATGAAGAAAGTTCCGGAGGCCCTATGAGAATGATGTGCTGCGCATG 947
Db 1024 GCCGAGCAGGAAATAATGAAGAAAGTTCCGGAGGCCCTATGAGAATGATGTGCTGCGCATG 1083
QY 948 AGC-----TCTCACTGATCCCTGGCTGAACCTGGCTGCTGCTAGTCTTTTCCAGCT 1001
Db 1084 AGCTGCGAGTCTCACTGATCCCTGGCTGAACCTGGCTGCTGCTAGTCTTTTCCAGCT 1143
QY 1002 TCATCAGGCGAGTCCCGCGCTCCAGCAGCGTTACTGGGCTGCTCCCTATAGCTCC 1061
Db 1144 TCATCAGGCGAGTCCCGCGCTCCAGCAGCGTTACTGGGCTGCTCCCTATAGCTCC 1203
QY 1062 TTTATCAGGCTCCCGAGCAGGAGATGGTGAGGTGTTATCCCGCCCGCAGGCGAGTGGC 1121
Db 1204 TTTATCAGGCTCCCGAGCAGGAGATGGTGAGGTGTTATCCCGCCCGCAGGCGAGTGGC 1263
QY 1122 GCCATCATCGGCAAGAGGGGACACATCAACAGCTCTCCCGGTTTGGCAGGCTCC 1181
Db 1264 GCCATCATCGGCAAGAGGGGACACATCAACAGCTCTCCCGGTTTGGCAGGCTCC 1323
QY 1182 ATCAAGATTGCACACCGGAAACACCTGACTCCAAAGTTGCTATGTTATCATCTGGA 1241
Db 1324 ATCAAGATTGCACACCGGAAACACCTGACTCCAAAGTTGCTATGTTATCATCTGGA 1383
QY 1242 CGCCAGAGGCCCAATTCAAGGCTCAGGAGAGATCTATGGCAAACTCAAGGAGGAGAAC 1301
Db 1384 CGCCAGAGGCCCAATTCAAGGCTCAGGAGAGATCTATGGCAAACTCAAGGAGGAGAAC 1443
QY 1302 TTTCTTTGTTCCCAAGAGGAGTGAAGTGGAGACCCACATACGTGTGCCAGCATCAGCA 1361
Db 1444 TTTCTTTGTTCCCAAGAGGAGTGAAGTGGAGACCCACATACGTGTGCCAGCATCAGCA 1503
QY 1362 GTTGGCGGGTCAATTGGCAAGGTGAAACCGTGAAACGAGTTGCAGAAATTTGACGGCA 1421
Db 1504 GTTGGCGGGTCAATTGGCAAGGTGAAACCGTGAAACGAGTTGCAGAAATTTGACGGCA 1563
QY 1422 GTTGAAGTGTGATACCAAGAGACCAAGCCCTGATGAGAAACCAAGGTCATCGTGAA 1481
Db 1564 GTTGAAGTGTGATACCAAGAGACCAAGCCCTGATGAGAAACCAAGGTCATCGTGAA 1623
QY 1482 ATCATCGGACATTTCTATGCTCAGTGGCTCAAGGAGATCCGAGACATCTGGCC 1541
Db 1624 ATCATCGGACATTTCTATGCTCAGTGGCTCAAGGAGATCCGAGACATCTGGCC 1683
QY 1542 CAGGTTAAGCAGCAGCATCAGAGGAGCAGAGTAACCAAGCCCGCAGGCGAGGAGAGTGA 1601
Db 1684 CAGGTTAAGCAGCAGCATCAGAGGAGCAGAGTAACCAAGCCCGCAGGCGAGGAGAGTGA 1743
QY 1602 CAGGCTTCCCTGCTGCTTNGAGTCCAGGACAAACCAAGGAGAAATCGAGAGTGTCT 1661
Db 1744 CAGGCTTCCCTGCTGCTTNGAGTCCAGGACAAACCAAGGAGAAATCGAGAGTGTCT 1803
QY 1662 CTCCTCCGCGAGGCTCAGAGATGATGGGAATCCGGGACACNTGGCGGCTCTAGATCA 1721
Db 1804 CTCCTCCGCGAGGCTCAGAGATGATGGGAATCCGGGACACNTGGCGGCTCTAGATCA 1863
QY 1722 GGTTCGCCCACTTGAATGAGAAAGATGTTCCAGTGAAGAACCTGATCTNTCAGGCCCAA 1781
Db 1864 GGTTCGCCCACTTGAATGAGAAAGATGTTCCAGTGAAGAACCTGATCTNTCAGGCCCAA 1923
QY 1782 ACACCCACCAATTTGGCCCAACACTGTNTTCCCTCTGGGGTGTGAGAAATTTAGCGCAA 1841
Db 1924 ACACCCACCAATTTGGCCCAACACTGTNTTCCCTCTGGGGTGTGAGAAATTTAGCGCAA 1983
```

```
QY 1842 GGCACCTTTAAACGTGATGTTTAAAGAGCTCTCCAGGCCCCACCAAGAGGTTGATC 1901
Db 1984 GGCACCTTTAAACGTGATGTTTAAAGAGCTCTCCAGGCCCCACCAAGAGGTTGATC 2043
QY 1902 ACACCTCAGTGGAGAGAAAAATAAATTTCTTCCAGTTTAAAA 1946
Db 2044 ACACCTCAGTGGAGAGAAAAATAAATTTCTTCCAGTTTAAAA 2088

RESULT 5
AF198254 2381 bp mRNA linear PRI 02-MAR-2000
LOCUS Homo sapiens mRNA-binding protein CRDBP mRNA, complete cds.
DEFINITION AF198254
ACCESSION AF198254.1 GI:7141071
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2381)
Ioannidis,P., Trangas,T., Dimitriadis,E., Samiotaki,M.,
Panoutsakopoulos,G., Kyriazoglou,I., Voutzoulis,S.,
Tsiapalis,C.M., Kittas,C., Agnantis,N. and Pandis,N.
Ectopic expression of a KH-domain containing protein, highly
homologous to both human IMP-1 and mouse CRD-BP, in benign and
malignant mesenchymal tumors
Unpublished
2 (bases 1 to 2381)
Ioannidis,P., Trangas,T., Dimitriadis,E. and Samiotaki,M.
Direct Submission
TITLE
Submitted (25-OCT-1999) G. Papanikolaou Research Center, St Savvas
Hospital, 171 Alexandras Avenue, Athens, Attica 11522, Greece
JOURNAL
FEATURES
Location/Qualifiers
1..2381
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q21.3-Q22"
/cell_line="MCF-7"
306..2039
/note="KH-domain containing protein; similar to Mus
musculus CRD-BP (c-myc mRNA coding region
determinant-binding protein) and Homo sapiens IMP-1
(insulin growth factor II mRNA-binding protein)"
/codon_start=1
/product="mRNA-binding protein CRDBP"
/protein_id="AAF37203.1"
/db_xref="GI:7141072"
/translation="MNKLYIGNESVTPADLEKVPFAEHKISYSGFLVKSGYAFVDC
PDEHMAKAIETFSKVELQGRLEIHSVPKQSRKQIRNIPQLRWELDSLILA
QYGTVENCEQVNTSETVAVVYTSNREQTRQAIKMLNGHLENHALKVSYTPDEQIA
QCPENRRGFGSGRQPGSPVAAGAPAKQOQVDIPLRLVPTQVVGAIIGKEGATI
RNIKTQSKIDVHRKENAGAAKAI SVHSTPEGSCACKMLEIMHKEAKTKTDAE
VPLKTLANNFVRLIKEGRNKKVEODTETKITISLQDLTLNPERTITVKGAIIE
NCRABEQLMKKRYREAYENDVAMSLQSHLTFLGLNLAAGLFPASSAVPPSPSVTG
AAPYSFWMQAEQWQVPIPAQVAGIIGKQHQIKQUSRFASASIKIAPETPDSK
VMVITITGPEAQKQRIYKGLKEENFFGPKREVKLETHIRVPASAGRVIGKGGK
TNEQLNLTAABVWVVRDQTPDENQIVIKIIFHYASQMAQRKIRDILAQVKQHQK
GQSNQAQARK"

Query Match 92.3%; Score 1796.6; DB 9; Length 2381;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1826; Conservative 0; Mismatches 9; Indels 7; Gaps 2;

QY 108 AGAGCGCGAAATTCAAATCCGAATATTCACCCAGCTCCGATGGGAGTACTGGAC 167
Db 540 AGAGCGCGAAATTCAAATCCGAATATTCACCCAGCTCCGATGGGAGTACTGGAC 599
QY 168 AGCCTCTCGCTCAGTATGTTACAGTAGAAGTGTGAGCAAGTGAACACCGAGAGTGAG 227
```

QY	1321	AAGTGAAGCTGGAGACCAACATACGTGTGCCAGCATCAGCAGCTGGCGCGGTCAATTGGCA	1380
Db	1321	AAGTGAAGCTGGAGACCAACATACGTGTGCCAGCATCAGCAGCTGGCGCGGTCAATTGGCA	1380
QY	1381	AAGTGAAAAACCGTGAAACAGATTGACAGATTTGACGGCAGCTGAGTGGTAGTACAA	1440
Db	1381	AAGTGAAAAACCGTGAAACAGATTGACAGATTTGACGGCAGCTGAGTGGTAGTACAA	1440
QY	1441	GAGACACAGACCCCTGATGAGAAACACAGAGTTCATCGTGAATAATCATCGGACATTTCTATG	1500
Db	1441	GAGACACAGACCCCTGATGAGAAACACAGAGTTCATCGTGAATAATCATCGGACATTTCTATG	1500
QY	1501	CCAGTCAGATGGCTCAACGGAAGATCCGAGACATCTCTGGCCAGGTTAAGCAGCAGCATC	1560
Db	1501	CCAGTCAGATGGCTCAACGGAAGATCCGAGACATCTCTGGCCAGGTTAAGCAGCAGCATC	1560
QY	1561	AGAGGACACAGAGTAAACCGGCCACGACGAGGAGAGTGCACAGCCCTCCCTGTCCT	1620
Db	1561	AGAGGACACAGAGTAAACCGGCCACGACGAGGAGAGTGCACAGCCCTCCCTGTCCT	1620
QY	1621	TNAGTCCAGGACAAACACGGCGAGAAATCGAGAGTGTCTCTCCCGGACAGGCTGAGA	1680
Db	1621	TNAGTCCAGGACAAACACGGCGAGAAATCGAGAGTGTCTCTCCCGGACAGGCTGAGA	1680
QY	1681	ATGAGTGGGAATCCGGGACACNTGGCGCGGCTGTAGATCAGTTTGCCTGATTTGA	1740
Db	1681	ATGAGTGGGAATCCGGGACACNTGGCGCGGCTGTAGATCAGTTTGCCTGATTTGA	1740
QY	1741	GAAGAGTGTTCAGTGAGGAAACCTGTATCTNTAGCCCAACACCCACCAATTTGGCCC	1800
Db	1741	GAAGAGTGTTCAGTGAGGAAACCTGTATCTNTAGCCCAACACCCACCAATTTGGCCC	1800
QY	1801	AACACTGNTGCCCTCGGGGTGTGAGAAATTTAGCCCAAGGCACATTTTAAAGCTGGAT	1860
Db	1801	AACACTGNTGCCCTCGGGGTGTGAGAAATTTAGCCCAAGGCACATTTTAAAGCTGGAT	1860
QY	1861	TGTTTAAAGAGCTCTCAGGCCCCCAAGAGAGGTGTGATCACACCTCAGTGGGAGAAA	1920
Db	1861	TGTTTAAAGAGCTCTCAGGCCCCCAAGAGAGGTGTGATCACACCTCAGTGGGAGAAA	1920
QY	1921	ATAAATAATTCCTTCAGTTTAAAA	1946
Db	1921	ATAAATAATTCCTTCAGTTTAAAA	1946
RESULT 4			
LOCUS	AF117106	2130 bp	mRNA
DEFINITION	Homo sapiens IGF-II mRNA-binding protein 1 (IMP-1) mRNA, complete cds.	linear	PRI 26-JAN-1999
ACCESSION	AF117106		
VERSION	AF117106.1	GI:4191607	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Nielsen, J., Christensen, J., Lykke-Andersen, J., Johnsen, A.H., Wewer, U.M. and Nielsen, F.C.		
TITLE	A family of insulin-like growth factor II mRNA-binding proteins represses translation in late development		
JOURNAL	Mol. Cell. Biol. 19 (2), 1262-1270 (1999)		
MEDLINE	99108099		
PUBMED	9891060		
REFERENCE			
AUTHORS	Nielsen, J., Christensen, J., Lykke-Andersen, J., Johnsen, A.H., Wewer, U.M. and Nielsen, F.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (30-DEC-1998) Institute of Molecular Biology, University of Copenhagen, Soelvgade 83H, Copenhagen DK-1307, Denmark		
FEATURES	Location/Qualifiers		

source	1. .2130	/organism="Homo sapiens"	
gene	/mol_type="mRNA"	/db_xref="taxon:9606"	
CDS	1. .2130	/gene="IMP-1"	
	10. .1743	/note="binds multiple sites in IGF-II mRNA 5'UTR"	
	/codon_start=1	/product="IGF-II mRNA-binding protein 1"	
	/protein_id="AAD09826.1"	/db_xref="GI:4191608"	
	/translation="MNKLYIGNLNESTVPADLEKVFPAEHKISYSGQELVKSGYAFVDC PDEHWAKAIETFSQVLEQRIEHSVPKORSKIQRNIPQURWLVLSLLA QYGTVENCEQVNTSETAVVNVTYSNREQRTQAIMKLNHGHLENHAKLVSVIPBQIA QGPNGRRGGFGSRGQSPVAGAPAKQQQVDIPLRLVPTQYVYGAIIQKEGATI RNIPTKOTSKI DVHRKENAGAAEKASVHSHPGCSACKMILIMKEAKDKTAD VPLKILAHNNFVGLIGKEGRNLKVBQDTETKITISSLODLTYLNPERTTVSGAIE NCCRAEQEIMKKVREAYENDVANSLSHLIPGLNLAAGLFPASSAVPPSPVTCG AAPYSSEMQPEQEMVQVFIQAQVAGALIGKRGQHIKQLSRFASASIKIAPETPDSK VRWVITITPPPEAQFKAQGRITYGKLEENFFPKPEVKLETHIRVPASAAGRVIGGGK TVNELQNLTAEEVVVPRDQTPDENDQIVIKIIGHFYASQMAQRKIRDILAQVKQHQK QGSNQAQARRK"		
ORIGIN			
Query Match	93.4%;	Score 1818;	DB 9; Length 2130;
Best Local Similarity	99.4%;	Pred. No. 0;	
Matches 1834;	Conservative 0;	Mismatches 5;	Indels 6; Gaps 1;
QY	108	AGGAGCGGAAAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTACTGGAC	167
Db	244	AGGAGCGGAAAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTACTGGAC	303
QY	168	AGCCTGTGCTCAGTATGTGTACAGTAGAGAACTGTGAGCAAGTGAAACACCGAGAGTGAG	227
Db	304	AGCCTGTGCTCAGTATGTGTACAGTAGAGAACTGTGAGCAAGTGAAACACCGAGAGTGAG	363
QY	228	ACGCGAGTGTGAATGTACCTATTCCAAACCGGAGCAGACACGAGCAAGCATCATGAAG	287
Db	364	ACGCGAGTGTGAATGTACCTATTCCAAACCGGAGCAGACACGAGCAAGCATCATGAAG	423
QY	288	CTGAATGGCCACCACTGGAGAACCATGCGCTGAAGGTCTCTCATATCCCGATGAGCAG	347
Db	424	CTGAATGGCCACCACTGGAGAACCATGCGCTGAAGGTCTCTCATATCCCGATGAGCAG	483
QY	348	ATAGCACAGGACCTGGAATGGCGCGCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGC	407
Db	484	ATAGCACAGGACCTGGAATGGCGCGCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGC	543
QY	408	CAGGGCTCACTGTGGCAGCGGGGGCCCCAGCCAGCAGCAGCAGCAGTGGACATCCCCCTT	467
Db	544	CAGGGCTCACTGTGGCAGCGGGGGCCCCAGCCAGCAGCAGCAGCAGTGGACATCCCCCTT	603
QY	468	CGGCTCTGTGGTCCCAACCCAGTATGTGGTGCCTATTATGGCAAGAGGGGGCCACCATC	527
Db	604	CGGCTCTGTGGTCCCAACCCAGTATGTGGTGCCTATTATGGCAAGAGGGGGCCACCATC	663
QY	528	CGCAACATCACAAAACAGACCCAGTCCAAAGTAGACGTGTCATAGGAAGAGAACGAGGT	587
Db	664	CGCAACATCACAAAACAGACCCAGTCCAAAGTAGACGTGTCATAGGAAGAGAACGAGGT	723
QY	588	GCAGCTGAAAAGGCATCAGTGTGTCATCTCCACCCCTGAGGGCTGTCTCTCGCTTGTAA	647
Db	724	GCAGCTGAAAAGGCATCAGTGTGTCATCTCCACCCCTGAGGGCTGTCTCTCGCTTGTAA	783
QY	648	ATGATCTTGGAGATTATGCATAAAGAGGCTAAGACACCAAAACCGGCTGACAGGTCTCC	707
Db	784	ATGATCTTGGAGATTATGCATAAAGAGGCTAAGACACCAAAACCGGCTGACAGGTCTCC	843
QY	708	CTGAAGATCTCTGGGCCCAATAAATCTTTAGGGGCTCTCATTTGGCAAGAGGACGGAAC	767
Db	844	CTGAAGATCTCTGGGCCCAATAAATCTTTAGGGGCTCTCATTTGGCAAGAGGACGGAAC	903

Db	1801	AAACCTGNTGCCCCCTCGGGGTCTCAGAAATTTNAGCGCAAGCATTTTAAACGTGGAT	1860	Db	181	AGTATGGTACAGTAGAGAACTGTGACCAAGTGAAACCCGAGAGTGAGACGGCAGTGGTGA	240
QY	1861	TGTTTAAAGAAGCTCTCCAGGCCCCCACCAGAGGGTGGATCACCACCTCAGTGGGAAGAAA	1920	QY	241	ATGTCACTTATTCACACCGGGAGCAGACACGAGCAAGCCATCATGAAGCTGAATGGCCACC	300
Db	1861	TGTTTAAAGAAGCTCTCCAGGCCCCCACCAGAGGGTGGATCACCACCTCAGTGGGAAGAAA	1920	Db	241	ATGTCACTTATTCACACCGGGAGCAGACACGAGCAAGCCATCATGAAGCTGAATGGCCACC	300
QY	1921	AATAAAATTCCTTCAGGTTTAAAA	1946	QY	301	AGTTGAGAACCATGCCCCCTGAAGGTCTCTCATATCCCGATGAGCAGATAGCAGGGAC	360
Db	1921	AATAAAATTCCTTCAGGTTTAAAA	1946	Db	301	AGTTGAGAACCATGCCCCCTGAAGGTCTCTCATATCCCGATGAGCAGATAGCAGGGAC	360
RESULT 3				QY	361	CTGAGAAATGGCGCGCAGAGGGGCTTTGGCTCTCGGGGTGAGCCGCCAGGGGTCACTCG	420
LOCUS	BD209926			Db	361	CTGAGAAATGGCGCGCAGAGGGGCTTTGGCTCTCGGGGTGAGCCGCCAGGGGTCACTCG	420
DEFINITION	Isolated nucleic acid molecules encoding cancer-associated			QY	421	TGSCAGCGGGGGCCCCCAGCACAAGCAGCAAGTGGACATCCCCCTTCCTGGCTCTCTGGTGC	480
ACCESSION	BD209926			Db	421	TGSCAGCGGGGGCCCCCAGCACAAGCAGCAAGTGGACATCCCCCTTCCTGGCTCTCTGGTGC	480
VERSION	BD209926.1	GI:33019696		QY	481	CCACCCAGTATGTGGGTGCGCATTTATGGCAAGAGGGGGCCACCATCCGCAACATCACAA	540
KEYWORDS	JP 2002512049-A/5.			Db	481	CCACCCAGTATGTGGGTGCGCATTTATGGCAAGAGGGGGCCACCATCCGCAACATCACAA	540
SOURCE	Homo sapiens (human)			QY	541	AAACAGACCCAGTCCCAAGATAGACGTGCATAGAAAGAGAGAACGACGTGACAGCTGAAAAAG	600
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			Db	541	AAACAGACCCAGTCCCAAGATAGACGTGCATAGAAAGAGAGAACGACGTGACAGCTGAAAAAG	600
REFERENCE	1 (bases 1 to 1946)			QY	601	CCATCAGTGTGCACTCCACCCCTGAGGGTGTCTCTCCCTGCTTGTAAAGATCATCTTGAGA	660
AUTHORS	Chen,Y.T., Gure,A., Tsang,S., Stockert,E., Jager,E., Knuth,A. and Old,J.J.			Db	601	CCATCAGTGTGCACTCCACCCCTGAGGGTGTCTCTCCCTGCTTGTAAAGATCATCTTGAGA	660
TITLE	Isolated nucleic acid molecules encoding cancer-associated			QY	661	TTATGATTAAGAGGCTAAGGACACCAAAACGGCTGACAGGTTCCCTTGAAGATCCTG	720
JOURNAL	Patent: JP 2002512049-A 5 23-APR-2002;			Db	661	TTATGATTAAGAGGCTAAGGACACCAAAACGGCTGACAGGTTCCCTTGAAGATCCTG	720
COMMENT	LUDWIG INSTITUTE FOR CANCER RESEARCH			QY	721	CCATAATAACTTTGTAGGGCGTCTCATTTGGCAAGAGGACCGAAACCTGGAAGAGTAG	780
	OS Homo sapiens (human)			Db	721	CCATAATAACTTTGTAGGGCGTCTCATTTGGCAAGAGGACCGAAACCTGGAAGAGTAG	780
	PN JP 2002512049-A/5			QY	781	AGCAAGATACCGAGACAAAATAACCATCTCTCTGTTGCAAGACCTTACCTTTTACAACC	840
	FF 16-MAR-1999 JP 2000545030			Db	781	AGCAAGATACCGAGACAAAATAACCATCTCTCTGTTGCAAGACCTTACCTTTTACAACC	840
	PR 17-APR-1998 US 09/061709			QY	841	CTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAAATTTGTTGCAGGGCCGACGAGAA	900
	PI YAO TSENG CHEN, ALI GURE, SOLAM TSANG, ELISABETH STOCKERT, ELKE			Db	841	CTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAAATTTGTTGCAGGGCCGACGAGAA	900
	PI JAGER			QY	901	TAATGAAGAAAGTTCCGGAGGGCTATGAGAATGATGTGGCTGCCATGAGCTCTCACCTGA	960
	PI ALEXANDER KNUTH, LLOYD J OLD			Db	901	TAATGAAGAAAGTTCCGGAGGGCTATGAGAATGATGTGGCTGCCATGAGCTCTCACCTGA	960
	PC C12N15/09,A61K35/12,A61K39/00,A61K39/39,A61P35/00,C07K16/32,			QY	961	TCCCTGGCTGAACTCTGGCTGTAGGTCTTTTCCAGCTTTCATCCAGGGCAGTCCCGC	1020
	PC C12N1/15,			Db	961	TCCCTGGCTGAACTCTGGCTGTAGGTCTTTTCCAGCTTTCATCCAGGGCAGTCCCGC	1020
	PC C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/53,G01N33/PC			QY	1021	CGCTCCAGCAGCGTTACTGGGGTGTCTTCCATAGCTCTTATGAGGCTCCCGAGC	1080
	574,			Db	1021	CGCTCCAGCAGCGTTACTGGGGTGTCTTCCATAGCTCTTATGAGGCTCCCGAGC	1080
	PC C12N15/00,C12N5/00			QY	1081	AGGAGATGCTGAGGTGTTTATCCCCCGCCAGGAGTGGCGCCATCATCGGCAAGAG	1140
	CC n=not determined(1622,1702,1771,1809,1833)			Db	1081	AGGAGATGCTGAGGTGTTTATCCCCCGCCAGGAGTGGCGCCATCATCGGCAAGAG	1140
	PH Key Location/Qualifiers			QY	1141	GGCAGCACATCAACAGCTCTCCGGTTTGCAGCGCCCTCCATCAAGATTGACACCCG	1200
	FT source 1..1946 /organism='Homo sapiens (human)'			Db	1141	GGCAGCACATCAACAGCTCTCCGGTTTGCAGCGCCCTCCATCAAGATTGACACCCG	1200
	FT Location/Qualifiers			QY	1201	AAACCTCAGTCCCAAGTTCGATGTTATCATCTGAGCGCCAGAGGAGGAGGAGG	1260
	1..1946			Db	1201	AAACCTCAGTCCCAAGTTCGATGTTATCATCTGAGCGCCAGAGGAGGAGGAGG	1260
	/organism='Homo sapiens'			QY	1261	AGGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1320
	/mol_type='genomic DNA'			Db	1261	AGGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1320
	/db_xref='taxon:9606'			QY			
ORIGIN				Db			
Query Match	99.7%; Score 1941; DB 6; Length 1946;			QY			
Best Local Similarity	100.0%; Pred. No. 0;			Db			
Matches 1946; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			QY	1	GCTGTAGCGAGGGGCTGGGGGCTGCTGTGCTCCCTTCTTCCGCGCTCGCGCTCAGC	60
				Db	1	GCTGTAGCGAGGGGCTGGGGGCTGCTGTGCTCCCTTCTTCCGCGCTCGCGCTCAGC	60
				QY	61	CCACCCAGAGCGGGGTGGAGGGGAGTGTCTGAGCTTCCCGGGTTAGGAGCCGAGAAA	120
				Db	61	CCACCCAGAGCGGGGTGGAGGGGAGTGTCTGAGCTTCCCGGGTTAGGAGCCGAGAAA	120
				QY	121	TTCAATCCGAATATTCACCCAGCTCCGATGGGAGTACTGGACAGGCTCTGCTC	180
				Db	121	TTCAATCCGAATATTCACCCAGCTCCGATGGGAGTACTGGACAGGCTCTGCTC	180
				QY	181	AGTATGTTACAGTAGAACTGTGAGCAAGTGAACCCGAGAGTGAACCGGACGTGGTGA	240

LOCUS AR343075 1946 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 7 from patent US 6576756.
ACCESSION AR343075
VERSION AR343075.1 GI:33738477
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1946)
AUTHORS Chen,Y.-T., Gure,A., Teang,S., Stockert,E., Jager,E., Alexander,K. and Old,L.J.
TITLE Isolated nucleic acid molecule encoding cancer associated antigen, the antigen itself, and uses thereof
JOURNAL Patent: US 6576756-A 7 10-JUN-2003;
FEATURES
source location/Qualifiers
1..1946
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 99.7%; Score 1941; DB 6; Length 1946;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1946; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCTGTAGCGAGGGGCTGGGGGCTGCTCTGTCCCTTCCTTGGCGCTGGCGCTCAGC 60
Db 1 GCTGTAGCGAGGGGCTGGGGGCTGCTCTGTCCCTTCCTTGGCGCTGGCGCTCAGC 60
Qy 61 CCACCCAGAGGCGGGGTGGGAGCGAGTCTCAGCTTCCCGGTTAGGAGCGGAAAA 120
Db 61 CCACCCAGAGGCGGGGTGGGAGCGAGTCTCAGCTTCCCGGTTAGGAGCGGAAAA 120
Qy 121 TTCAATCCGAAATATTCCACCCAGCTCCGATGGGAAGTACTGGACAGCCTGTGGCTC 180
Db 121 TTCAATCCGAAATATTCCACCCAGCTCCGATGGGAAGTACTGGACAGCCTGTGGCTC 180
Qy 181 AGTATGGTACGTAGAGACTGTGAGCAAGTGAACACCGAGTGAAGCGAGTGGTGA 240
Db 181 AGTATGGTACGTAGAGACTGTGAGCAAGTGAACACCGAGTGAAGCGAGTGGTGA 240
Qy 241 ATGTCACCTATTCCAAACGGGAGCGACACGAGCGACCATCATGAAGCTGAATGCCACC 300
Db 241 ATGTCACCTATTCCAAACGGGAGCGACACGAGCGACCATCATGAAGCTGAATGCCACC 300
Qy 301 AGTTGGAGAACCATGCCCTGAAGTCTCTTACATCCCGATGAGCAGATAGCACAGGAC 360
Db 301 AGTTGGAGAACCATGCCCTGAAGTCTCTTACATCCCGATGAGCAGATAGCACAGGAC 360
Qy 361 CTGAGATGGGCGCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGCAGGCTCACCTG 420
Db 361 CTGAGATGGGCGCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGCAGGCTCACCTG 420
Qy 421 TGGCAGCGGGGCCCCAGCAGCAGCAAGTGGACATCCCGCTTCGGCTCCTGGTGC 480
Db 421 TGGCAGCGGGGCCCCAGCAGCAGCAAGTGGACATCCCGCTTCGGCTCCTGGTGC 480
Qy 481 CCACCCAGTATGTGGTGCCATTATTTGCCAAGGAGGGGCCACCATCCGCAACATCAAA 540
Db 481 CCACCCAGTATGTGGTGCCATTATTTGCCAAGGAGGGGCCACCATCCGCAACATCAAA 540
Qy 541 AACAGACCCAGTCCAGATAGACGTGCATAGGAGAGAACCGCAGTGCAGTGAARAAG 600
Db 541 AACAGACCCAGTCCAGATAGACGTGCATAGGAGAGAACCGCAGTGCAGTGAARAAG 600
Qy 601 CCATCAGTGTGCATCCACCCCTGAGGGTGTCTCTCGCTTGTAAAGTATCTTGGAGA 660
Db 601 CCATCAGTGTGCATCCACCCCTGAGGGTGTCTCTCGCTTGTAAAGTATCTTGGAGA 660
Qy 661 TTATGCAATAAGAGCTTAAGACACCAAAAACGGCTGACAGAGTTCCCTTGAAGATCTCTGG 720
Db 661 TTATGCAATAAGAGCTTAAGACACCAAAAACGGCTGACAGAGTTCCCTTGAAGATCTCTGG 720
Qy 721 CCCATAATACTTTGTAGGGCGTCTCATTTGGCAAGGAGGAGCGAACCTGAAGAAGTAG 780

Db 721 CCCATAATACTTTGTAGGGCGTCTCATTTGCAAGGAGGAGGACCTGAAGAAGTAG 780
Qy 781 AGCAAGATACCGAGACAAAAATCCATCTCTCTGTTGCAAGACCTTACCTTTTCAACC 840
Db 781 AGCAAGATACCGAGACAAAAATCCATCTCTCTGTTGCAAGACCTTACCTTTTCAACC 840
Qy 841 CTGAGAGGACCATCACTGTGAAGGGGCCATCGAGAATTGTGAGGCGCGAGGAGAAA 900
Db 841 CTGAGAGGACCATCACTGTGAAGGGGCCATCGAGAATTGTGAGGCGCGAGGAGAAA 900
Qy 901 TAATGAAGAAAGTTCCGGAGGCCCTATGAGAAATGATGTGGCTGCCATGAGCTCTCACCTGA 960
Db 901 TAATGAAGAAAGTTCCGGAGGCCCTATGAGAAATGATGTGGCTGCCATGAGCTCTCACCTGA 960
Qy 961 TCCCTGGCTGAACCTGGCTCTGTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGC 1020
Db 961 TCCCTGGCTGAACCTGGCTCTGTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGC 1020
Qy 1021 CGCCTCCAGCAGCGTTACTTGGGCTGCTCCCTATAGCTCCTTTATGAGGCTCCCGAGC 1080
Db 1021 CGCCTCCAGCAGCGTTACTTGGGCTGCTCCCTATAGCTCCTTTATGAGGCTCCCGAGC 1080
Qy 1081 AGGAGATGCTCAGGTGTTTATCCCGCCAGGCGAGTGGCGGCCATCATCGGCAAGAAG 1140
Db 1081 AGGAGATGCTCAGGTGTTTATCCCGCCAGGCGAGTGGCGGCCATCATCGGCAAGAAG 1140
Qy 1141 GGCAGCAGATCAAAACAGCTCTCCCGGTTTGGCAGCGCTCCATCAAGATTGCACACCG 1200
Db 1141 GGCAGCAGATCAAAACAGCTCTCCCGGTTTGGCAGCGCTCCATCAAGATTGCACACCG 1200
Qy 1201 AAACACCTGACTCCAAAGTTCTGTTATGTTATCATCTGAGCGCGCAGAGGCCCAATTCA 1260
Db 1201 AAACACCTGACTCCAAAGTTCTGTTATGTTATCATCTGAGCGCGCAGAGGCCCAATTCA 1260
Qy 1261 AGGCTCAGGGAAGAACTTATGCGAAACTCAAGGAGGAGAACTTTTGGTCCCAAGAGG 1320
Db 1261 AGGCTCAGGGAAGAACTTATGCGAAACTCAAGGAGGAGAACTTTTGGTCCCAAGAGG 1320
Qy 1321 AAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGTGGCGGGTCAATTGGA 1380
Db 1321 AAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGTGGCGGGTCAATTGGA 1380
Qy 1381 AAGTGGAAAAACCGTGAACCGAGTTGAGAAATTTGACGCGAGCTGAGTGGTAGTACCAA 1440
Db 1381 AAGTGGAAAAACCGTGAACCGAGTTGAGAAATTTGACGCGAGCTGAGTGGTAGTACCAA 1440
Qy 1441 GAGACCAACCCCTGATGAGAAACGACAGGTCTATCGTGAATAATCATCGACATTTCTATG 1500
Db 1441 GAGACCAACCCCTGATGAGAAACGACAGGTCTATCGTGAATAATCATCGACATTTCTATG 1500
Qy 1501 CCAGTCAGATGGCTCAACGGAAGATCCGAGACATCTCTGGCCGAGTTAAGCAGCAGCATC 1560
Db 1501 CCAGTCAGATGGCTCAACGGAAGATCCGAGACATCTCTGGCCGAGTTAAGCAGCAGCATC 1560
Qy 1561 AGAGGGACAGAGTAAACAGGCCCGAGCAGCGAGGAAGTGAACCGCCCTCCTCTCCCT 1620
Db 1561 AGAGGGACAGAGTAAACAGGCCCGAGCAGCGAGGAAGTGAACCGCCCTCCTCTCCCT 1620
Qy 1621 TNGAGTCCAGACAAACACGCGCAGAAATCGAGAGTGTGCTCTCCCGCGAGCGCTGAGA 1680
Db 1621 TNGAGTCCAGACAAACACGCGCAGAAATCGAGAGTGTGCTCTCCCGCGAGCGCTGAGA 1680
Qy 1681 ATGAGTGGGAATCCGGGACACNTGGGCGGGGTGTAGATCAGGTTTGGCCATTGATGA 1740
Db 1681 ATGAGTGGGAATCCGGGACACACNTGGGCGGGGTGTAGATCAGGTTTGGCCATTGATGA 1740
Qy 1741 GAAAGATCTCCAGTGAAGAACCCCTGATCTNTCAGCCCAACACCCACCCAAATTTGGCC 1800
Db 1741 GAAAGATCTCCAGTGAAGAACCCCTGATCTNTCAGCCCAACACCCACCCAAATTTGGCC 1800
Qy 1801 AACACTGNTGCCCTCGGGGTGTGAGAAATTTAGCGCAAGGCACTTTTAAACGTGGAT 1860

JOURNAL Patent: US 5297364-A 7 02-OCT-2001;
 FEATURES location/Qualifiers
 source 1..1946
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN

Query Match 99.7%; Score 1941; DB 6; Length 1946;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1946; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTGTACGAGGGGCTGGGGGCTGCTGTGTCCTTCCCTGCGCTGCGGCTCAGC 60
 DB 1 GCTGTACGAGGGGCTGGGGGCTGCTGTGTCCTTCCCTGCGCTGCGGCTCAGC 60

QY 61 CCACCCAGAGGCCGGGTGGAGGGCGAGTCTCAGCTTCCCGGTTAGGAGCCGGA 120
 DB 61 CCACCCAGAGGCCGGGTGGAGGGCGAGTCTCAGCTTCCCGGTTAGGAGCCGGA 120

QY 121 TTCAATCCGAATATTTCCACCCAGCTCCGATGGGAAGTACTGGACAGCTGCTGCTC 180
 DB 121 TTCAATCCGAATATTTCCACCCAGCTCCGATGGGAAGTACTGGACAGCTGCTGCTC 180

QY 181 AGTATGTTACAGTAGAGAACTGTGACCAAGTGAACCCGAGAGTGAGACGCGAGTGTGA 240
 DB 181 AGTATGTTACAGTAGAGAACTGTGACCAAGTGAACCCGAGAGTGAGACGCGAGTGTGA 240

QY 241 ATGTACCTATTCACCCGGAGCAGACCGAGCAAGCCATCATGAAGCTGAATGGCCACC 300
 DB 241 ATGTACCTATTCACCCGGAGCAGACCGAGCAAGCCATCATGAAGCTGAATGGCCACC 300

QY 301 AGTGGAGACCATGCTCCCTGAAGTCTCCTACATCCCGGATGAGCAGATGACAGGGAC 360
 DB 301 AGTGGAGACCATGCTCCCTGAAGTCTCCTACATCCCGGATGAGCAGATGACAGGGAC 360

QY 361 CTGAGATGGCGCGCAGGGGGCTTTGGCTCTCGGGGTGAGCCCGCAGGGCTCACTG 420
 DB 361 CTGAGATGGCGCGCAGGGGGCTTTGGCTCTCGGGGTGAGCCCGCAGGGCTCACTG 420

QY 421 TGGCAGGGGGCCCCAGCAGCAGCAAGTGGACATCCCTTCCTGGTGC 480
 DB 421 TGGCAGGGGGCCCCAGCAGCAGCAAGTGGACATCCCTTCCTGGTGC 480

QY 481 CCACCCAGTATGGGTGCGCATTTATGGCAAGAGGGGGCCACCATCCGCAACATCA 540
 DB 481 CCACCCAGTATGGGTGCGCATTTATGGCAAGAGGGGGCCACCATCCGCAACATCA 540

QY 541 AACAGACCCAGTCCAGATAGACGTGATAGGAGGAGAACGCGAGGTGCAGCTGAA 600
 DB 541 AACAGACCCAGTCCAGATAGACGTGATAGGAGGAGAACGCGAGGTGCAGCTGAA 600

QY 601 CCATCAGTGTGACTCCACCCCTGAGGGCTGCTCCTCGGTGTAAGATGATCTTGAGA 660
 DB 601 CCATCAGTGTGACTCCACCCCTGAGGGCTGCTCCTCGGTGTAAGATGATCTTGAGA 660

QY 661 TTATGATGAAGAGGCTAAGGACCAAAAGCGCTGACAGGTTCCTCTGAAGTCTG 720
 DB 661 TTATGATGAAGAGGCTAAGGACCAAAAGCGCTGACAGGTTCCTCTGAAGTCTG 720

QY 721 CCATATATATCTTTAGGGCGTCTCATTTGGCAAGGAGGACCGAACTCGAAGGTAG 780
 DB 721 CCATATATATCTTTAGGGCGTCTCATTTGGCAAGGAGGACCGAACTCGAAGGTAG 780

QY 781 AGCAAGTACCGAGACAAAATACCATCTCTCGTTGCGAAGACCTTACCTTTACACC 840
 DB 781 AGCAAGTACCGAGACAAAATACCATCTCTCTGTTGCAAGACCTTACCTTTACACC 840

QY 841 CTGAGGACCATCATCTGAGGGGCGCATCGAGATTGTTGACGGGCGCAGCAGAAA 900
 DB 841 CTGAGGACCATCATCTGAGGGGCGCATCGAGATTGTTGACGGGCGCAGCAGAAA 900

QY 901 TAATGAAGAAAGTTCCGGAGGCGCTATGAGATGATGTGGCTGCCATGAGCTCTCACCTGA 960
 DB 901 TAATGAAGAAAGTTCCGGAGGCGCTATGAGATGATGTGGCTGCCATGAGCTCTCACCTGA 960

DB 901 TAATGAAGAAAGTTCCGGAGGCGCTATGAGATGATGTGGCTGCCATGAGCTCTCACCTGA 960
 QY 961 TCCTTGGCTGAAACCTGGCTGCTGTAGGTCTTTTCCAGCTTTCATCCAGGCGAGTCCCGC 1020
 DB 961 TCCTTGGCTGAAACCTGGCTGCTGTAGGTCTTTTCCAGCTTTCATCCAGGCGAGTCCCGC 1020

QY 1021 CGCTTCCAGCAGCGTTTACTTGGGGCTGCTCCCTATAGCTCTCTTATGACGCTCCGAGC 1080
 DB 1021 CGCTTCCAGCAGCGTTTACTTGGGGCTGCTCCCTATAGCTCTCTTATGACGCTCCGAGC 1080

QY 1081 AGGAGATGTTGTCAGGTGTTTATCCCGCCAGCAGTGGCGCCATCATCGCAAGAAG 1140
 DB 1081 AGGAGATGTTGTCAGGTGTTTATCCCGCCAGCAGTGGCGCCATCATCGCAAGAAG 1140

QY 1141 GGAGCAGCATCAAAACAGCTCTCCCGGTTTGGCAGCGCTCCATCAAGATTGCAACCCG 1200
 DB 1141 GGAGCAGCATCAAAACAGCTCTCCCGGTTTGGCAGCGCTCCATCAAGATTGCAACCCG 1200

QY 1201 AAACACTGACTCAAAAGTTCGTATGGTTATCATCTAGACCGCCAGAGGCCAATCA 1260
 DB 1201 AAACACTGACTCAAAAGTTCGTATGGTTATCATCTAGACCGCCAGAGGCCAATCA 1260

QY 1261 AGCTCAGGGAAGAACTCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGAGG 1320
 DB 1261 AGCTCAGGGAAGAACTCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGAGG 1320

QY 1321 AAGTGAAGCTGGAGACCCACATACGTGTGCAGCATCAGCAGCTGCGGGTCAATTGGCA 1380
 DB 1321 AAGTGAAGCTGGAGACCCACATACGTGTGCAGCATCAGCAGCTGCGGGTCAATTGGCA 1380

QY 1381 AAGTGGAAAAACGGTGAACGAGTTCAGAAATTTGACGAGCTGAGGTGGTAGTACCAA 1440
 DB 1381 AAGTGGAAAAACGGTGAACGAGTTCAGAAATTTGACGAGCTGAGGTGGTAGTACCAA 1440

QY 1441 GAGACAGACCCCTGATGAGAACAGCAGCATCTGTGAAAAATCATCGGACATTTCTATG 1500
 DB 1441 GAGACAGACCCCTGATGAGAACAGCAGCATCTGTGAAAAATCATCGGACATTTCTATG 1500

QY 1501 CCAGTCAAGTGGCTCAACCGAAGATCCGAGACATCTCGGCCAGGTTAAGCAGCAGCATC 1560
 DB 1501 CCAGTCAAGTGGCTCAACCGAAGATCCGAGACATCTCGGCCAGGTTAAGCAGCAGCATC 1560

QY 1561 AGAAGGACAGAGTAAACCGGCGCAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1620
 DB 1561 AGAAGGACAGAGTAAACCGGCGCAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1620

QY 1621 TNGAGTCCAGGACAAACAAACCGGCGCAGAAATCGAGAGTGTCTCTCCCGGAGGCTGAGA 1680
 DB 1621 TNGAGTCCAGGACAAACAAACCGGCGCAGAAATCGAGAGTGTCTCTCCCGGAGGCTGAGA 1680

QY 1681 ATGAGTGGGAATCCGGGACACNTGGGCGGGCTGTAGATCAGGTTTGGCCCATTTGATTGA 1740
 DB 1681 ATGAGTGGGAATCCGGGACACNTGGGCGGGCTGTAGATCAGGTTTGGCCCATTTGATTGA 1740

QY 1741 GAAAGATGTTCCAGTCAAGGAACCTCATCTNTCAGCCCAACACCCCAACCAATTTGGCCC 1800
 DB 1741 GAAAGATGTTCCAGTCAAGGAACCTCATCTNTCAGCCCAACACCCCAACCAATTTGGCCC 1800

QY 1801 AACACTGTNTGCCCTCGGGGTGTCAAAATTTNAGCGAAGGCACTTTTAAACGTTGGAT 1860
 DB 1801 AACACTGTNTGCCCTCGGGGTGTCAAAATTTNAGCGAAGGCACTTTTAAACGTTGGAT 1860

QY 1861 TGTTTAAAGAGCTCTCCAGGGCCCAACCAAGAGGGTGGATCACTCAGTGGGAAGAAA 1920
 DB 1861 TGTTTAAAGAGCTCTCCAGGGCCCAACCAAGAGGGTGGATCACTCAGTGGGAAGAAA 1920

QY 1921 AATAAAATTTCTTCAGGTTTAAAA 1946
 DB 1921 AATAAAATTTCTTCAGGTTTAAAA 1946

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2004, 03:38:17 ; Search time 7493.35 Seconds
(without alignments)
11256.058 Million cell updates/sec

Title: US-09-270-437D-7
Perfect score: 1946
Sequence: 1 gctgtacgaggggctggg.....atttccttcgggttttaaaa 1946

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_on.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vl.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rtd.*
36: em_htg_nam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1941	99.7	1946	6	AR171866	AR171866 Sequence
2	1941	99.7	1946	6	AR343075	AR343075 Sequence
3	1941	99.7	1946	6	BD209926	BD209926 Isolated
4	1818	93.4	2130	9	AF117106	AF117106 Homo sapi
5	1796.6	92.3	2381	9	AF198254	AF198254 Homo sapi
6	1669.2	85.8	1708	6	AR171864	AR171864 Sequence
7	1669.2	85.8	1708	6	AR343073	AR343073 Sequence
8	1669.2	85.8	1708	6	BD209924	BD209924 Isolated
9	1389	71.4	2780	6	BD127811	BD127811 Primer fo
10	1389	71.4	2780	9	AK074915	AK074915 Homo sapi
11	1301	66.9	2224	6	AR160244	AR160244 Sequence
12	1298.6	66.7	2223	10	AF061569	AF061569 Mus muscu
13	1297.6	66.7	2444	10	BC051679	BC051679 Mus muscu
14	1250	64.2	2105	10	AF541940	AF541940 Rattus no
15	1046	53.8	2021	5	AF026527	AF026527 Gallus ga
16	732.8	37.7	833	6	BD124998	BD124998 Primer fo
17	732.8	37.7	833	6	BD126684	BD126684 Primer fo
18	723.4	37.2	1740	10	AB046173	AB046173 Mus muscu
19	723.4	37.2	2140	10	BC045138	BC045138 Mus muscu
20	723.4	37.2	4030	10	BC049082	BC049082 Mus muscu
21	674.2	34.6	1740	6	BD275302	BD275302 COMPOUNDS
22	674.2	34.6	1740	6	AR220687	AR220687 Sequence
23	674.2	34.6	1740	6	AR281251	AR281251 Sequence
24	674.2	34.6	1740	6	AX365954	AX365954 Sequence
25	674.2	34.6	1743	6	AX366054	AX366054 Sequence
26	674.2	34.6	1743	6	AX366057	AX366057 Sequence
27	674.2	34.6	1764	6	AX366035	AX366035 Sequence
28	674.2	34.6	4171	9	AF117108	AF117108 Homo sapi
29	672.6	34.6	4155	9	HSU76705	U76705 Human putat
30	672.6	34.6	4159	6	AR171863	AR171863 Sequence
31	672.6	34.6	4159	6	AR343072	AR343072 Sequence
32	672.6	34.6	4159	6	BD209923	BD209923 Isolated
33	672.6	34.6	4181	6	BD275767	BD275767 COMPOUNDS
34	672.6	34.6	4181	6	AR220552	AR220552 Sequence
35	672.6	34.6	4181	6	AR255546	AR255546 Sequence
36	672.6	34.6	4181	6	AR281116	AR281116 Sequence
37	672.6	34.6	4181	6	AX333233	AX333233 Sequence
38	672.6	34.6	4181	6	AX365782	AX365782 Sequence
39	672.6	34.6	4181	9	HSU97188	U97188 Homo sapien
40	672.6	34.6	4602	6	AX397963	AX397963 Sequence
41	667.6	34.3	2011	5	AF161270	AF161270 Danio rer
42	667.6	34.3	2612	5	BC045873	BC045873 Danio rer
43	661.4	34.0	2161	5	AF055923	AF055923 Xenopus l
44	661.4	34.0	2405	5	AF064634	AF064634 Xenopus l
45	650.8	33.4	2031	5	BC057700	BC057700 Xenopus l

ALIGNMENTS

RESULT 1
LOCUS AR171866
DEFINITION Sequence 7 from patent US 6297364.
ACCESSION AR171866
VERSION AR171866.1 GI:17910816
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1946)
AUTHORS Chen,Y.-T., Gure,A., Tsang,S., Stockert,E., Jager,E., Alexander,K.
and Old,L.J.
TITLE Isolated nucleic acid molecule encoding cancer associated antigen,
the antigen itself, and uses thereof

1946 bp DNA linear PAT 17-DEC-2001